

JAP 10

Search statement 18

?

** SS 18: Results 3

Search statement 19

?

1/3 JAPIO - (C) JPO- image

PN - JP 2000078577 A 20000314 [JP2000078577]
TI - METHOD AND SYSTEM FOR PROCESSING MULTIPLEX STREAM OF VIDEO FRAME
IN - LIRA BOROKUSKI; AGNES Y NYE; EDWARD F WESTERMAN
PA - INTERNATL BUSINESS MACH CORP <IBM>
AP - JP14634199 19990526 [1999JP-0146341]
PR - US98 87603 19980529 [1998US-0087603]
IC1 - H04N-007/24
IC2 - G06T-001/20 H04J-003/00 H04N-007/08 H04N-007/081
AB - PROBLEM TO BE SOLVED: To obtain multiplex program compression technique for optimizing video quality in a whole system by dynamically adapting the encoding of a multiplex stream in a video frame based on the correlative complex degree of the video frame constituting the multiplex stream of the video frame.
- SOLUTION: Respective encoders 210 generates statistics concerning a picture which is just encoded. The statistics are made to be an input parameter to a connection rate control algorithm 230 and the algorithm 230 dynamically calculates the bit rate of each encoder 210. The bit rate of each encoder 210 is decoded based on the correlative complex degree of programs (source 1-source n) and the occurrence of scene change in the corresponding program. Each encoder 210 generates an interchangeable bit stream in MPEG-2 standard. Each encoder 210 changes the bit rate in a GOP boundary or scene change in accordance with a connection rate controller 230.
- COPYRIGHT: (C)2000,JPO

2/3 JAPIO - (C) JPO- image

PN - JP 63286083 A 19881122 [JP63286083]
TI - DEVICE FOR BUFFERING DIGITAL PICTURE SIGNAL
IN - NAGAI MICHIO; KONDO TETSUJIRO
PA - SONY CORP
AP - JP12162487 19870519 [1987JP-0121624]
IC1 - H04N-007/13
IC2 - G06F-015/66 H04B-014/04 H04N-005/92
AB - PURPOSE: To contrive to reduce the memory capacity and quantization distortion by using plural threshold values offering to decrease the degree of quantization distortion, compressing a picture element data of a block by a bit number depending on the dynamic range for each block and smaller than the original quantization bit number and coding the result.
- CONSTITUTION: A high efficiency code (ADRC) coding is applied to an inputted digital video signal by the ADRC encoder 2 and a dynamic range DR is detected. The DR is fed to a threshold value generating section 18 and an integrated frequency distribution is obtained.

BEST AVAILABLE COPY

1/3 JAPIO - (C) JPO- image

PN - JP 2000078577 A 20000314 [JP2000078577]

TI - METHOD AND SYSTEM FOR PROCESSING MULTIPLEX STREAM OF VIDEO FRAME

IN - LIRA BOROKUSKI; AGNES Y NYE; EDWARD F WESTERMAN

PA - INTERNATL BUSINESS MACH CORP <IBM>

AP - JP14634199 19990526 [1999JP-0146341]

PR - US98 87603 19980529 [1998US-0087603]

IC1 - H04N-007/24

IC2 - G06T-001/20 H04J-003/00 H04N-007/08 H04N-007/081

AB - PROBLEM TO BE SOLVED: To obtain multiplex program compression technique for optimizing video quality in a whole system by dynamically adapting the encoding of a multiplex stream in a video frame based on the correlative complex degree of the video frame constituting the multiplex stream of the video frame.

- SOLUTION: Respective encoders 210 generates statistics concerning a picture which is just encoded. The statistics are made to be an input parameter to a connection rate control algorithm 230 and the algorithm 230 dynamically calculates the bit rate of each encoder 210. The bit rate of each encoder 210 is decoded based on the correlative complex degree of programs (source 1-source n) and the occurrence of scene change in the corresponding program. Each encoder 210 generates an interchangeable bit stream in MPEG-2 standard. Each encoder 210 changes the bit rate in a GOP boundary or scene change in accordance with a connection rate controller 230.

- COPYRIGHT: (C)2000,JPO

Search statement 19

?

Subaccount END800-0161

JAPIO - Time in minutes : 2,04

The cost estimation below is based on Questel's standard price list

Estimated cost :	3.40 USD
Records displayed and billed :	4
Estimated cost :	6.00 USD
Cost estimated for the last database search :	9.40 USD
Estimated total session cost :	64.55 USD

Selected file: INSPEC

Compiled and produced by the IEE in association with Fiz Karlsruhe

Copyright 1998 Institution of Electrical Engineers (IEE)

The document delivery service is available with ..OR command

Covers from 1969 to : 2001-02-19.

Search statement 1

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-78577

(P2000-78577A)

(43) 公開日 平成12年3月14日 (2000.3.14)

(51) Int.Cl. ⁷	識別記号	F I	テマコード* (参考)
H 0 4 N	7/24	H 0 4 N 7/13	Z
G 0 6 T	1/20	H 0 4 J 3/00	M
H 0 4 J	3/00	G 0 6 F 15/66	K
H 0 4 N	7/08	H 0 4 N 7/08	Z
	7/081		

審査請求 有 請求項の数60 OL (全 27 頁)

(21) 出願番号	特願平11-146341	(71) 出願人	390009531 インターナショナル・ビジネス・マシーンズ・コーポレーション INTERNATIONAL BUSINESS MACHINES CORPORATION アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
(22) 出願日	平成11年5月26日 (1999.5.26)	(74) 代理人	100086243 弁理士 坂口 博 (外1名)
(31) 優先権主張番号	09/087603		
(32) 優先日	平成10年5月29日 (1998.5.29)		
(33) 優先権主張国	米国 (US)		

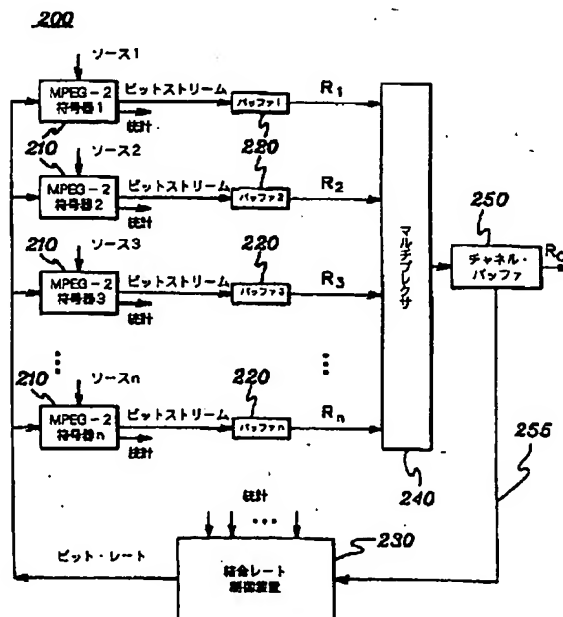
最終頁に続く

(54) 【発明の名称】 ビデオ・フレームの多重ストリームを処理する方法及びシステム

(57) 【要約】 (修正有)

【課題】 使用可能な帯域幅を番組間で動的に分配し、システムの全体的なビデオ品質を最適化する多重番組圧縮技術を提供する。

【解決手段】 ビデオ・データの多重ストリームを並列に動的に符号化し、一定ビット・レート・チャンネル上に多重化する場合、各ビデオ・データ・ストリームに対する個々の符号化ビット・レートが、ビデオ・データの多重ストリームの相対複雑度、並びに圧縮ビデオ・データ・バッファ及び符号器と一定ビット・レート・チャンネル間に接続されるチャンネル・バッファの充填度に部分的に基づき、ビデオ・フレームの多重ストリームを並列に符号化するステップと、ビデオ・フレームの相対複雑度に基づき、ビデオ・フレームの少なくとも1つのストリームの符号化を動的に適応化するステップとを含む。各ストリームのビット・レートは、GOP境界においてだけ、又はシーン変化が発生する場合、更にバッファ充填度にもとづき変更される。



【特許請求の範囲】

【請求項1】ビデオ・フレームの多重ストリームを処理する方法であって、

前記ビデオ・フレームの多重ストリームを並列に符号化するステップと、

前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、少なくとも1つのストリームの符号化を動的に適応化するステップとを含む、方法。

【請求項2】前記少なくとも1つのストリームの符号化を動的に適応化するステップを、当該ストリーム内のシーン変化の検出に際して実行するステップを含む、請求項1記載の方法。

【請求項3】前記符号化を動的に適応化するステップが、前記少なくとも1つのストリーム内の前記シーン変化の検出に際して、当該ストリームを符号化するために使用される制御可能な符号化パラメータを調整するステップを含む、請求項2記載の方法。

【請求項4】前記少なくとも1つのストリーム内の前記シーン変化を検出するステップを含む、請求項3記載の方法。

【請求項5】前記符号化を動的に適応化するステップが、前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、前記多重ストリームの各ストリームの符号化を動的に適応化するステップを含む、請求項1記載の方法。

【請求項6】前記符号化を動的に適応化するステップが、前記符号化により使用される制御可能な符号化パラメータを調整し、前記各ストリームを符号化するステップを含む、請求項5記載の方法。

【請求項7】前記各ストリームに対して、前記制御可能な符号化パラメータを調整する前記ステップが、当該ストリーム内のシーン変化の検出時に、または新たなフレーム・グループ（GOP）の開始時に発生する、請求項6記載の方法。

【請求項8】前記ストリーム内のシーン変化の検出無しに、新たなフレーム・グループが開始するとき、前記制御可能な符号化パラメータの調整を、所定の割合調整に制限するステップを含む、請求項7記載の方法。

【請求項9】シーン変化の検出に際して、前記符号化を動的に適応化するステップが、現フレーム・グループ（GOP）を終了し、新たなGOPを開始するステップを含み、前記制御可能な符号化パラメータを調整するステップが、前記所定の割合調整を使用することなく、前記制御可能な符号化パラメータを調整するステップを含む、請求項8記載の方法。

【請求項10】前記多重ストリームの隣接ビデオ・フレームをモニタすることにより、前記シーン変化を検出するステップを含む、請求項9記載の方法。

【請求項11】前記各ストリームに対する前記制御可能な符号化パラメータが、符号化ビット・レートを含む、

請求項7記載の方法。

【請求項12】前記符号化が前記各ストリームに対して、当該ストリーム内のシーン変化の検出時に、新たなフレーム・グループ（GOP）を開始するステップを含み、前記各ストリームに対して、前記符号化ビット・レートが各GOP内で一定であり、前記GOP間で可変である、請求項11記載の方法。

【請求項13】前記符号化が多重圧縮ビデオ・ストリームを生成し、前記方法が、各圧縮ビデオ・ストリームをバッファ内にバッファリングするステップを含み、前記符号化を動的に適応化するステップが、前記各ストリームに対して、前記圧縮ビデオ・ストリームを受信する前記バッファの充填度にもとづき、前記符号化ビット・レートを変更するステップを含む、請求項11記載の方法。

【請求項14】前記符号化ビット・レートを変更するステップが、前記符号化ビット・レートを $R_{i,old} = R_i + E \cdot f_i / N_i$ として変更するステップを含み、ここで R_i はビデオ・フレームのストリーム i に対して計算されたビット・レート、 E はバッファ充填度誤差ビット、 N_i はストリーム i のGOP内のピクチャ数、 f_i はストリーム i のフレーム・レートである、請求項13記載の方法。

【請求項15】前記多重圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化するステップを含む、請求項13記載の方法。

【請求項16】前記多重圧縮ビデオ・ストリームを、前記一定ビット・レート・チャンネルに接続されるチャンネル・バッファ内にバッファリングするステップを含み、前記符号化を動的に適応化するステップが、前記各ストリームに対して、前記チャンネル・バッファの充填度にもとづき、前記符号化ビット・レートを変更するステップを含む、請求項15記載の方法。

【請求項17】前記チャンネル・バッファの上限及び下限において、保護帯域を事前に定義するステップを含み、前記バッファ充填度が前記チャンネル・バッファの前記上下のいずれかの保護帯域内にある場合、前記符号化ビット・レートを変更するステップを含む、請求項16記載の方法。

【請求項18】前記符号化ビット・レートを変更するステップが、前記バッファ充填度が前記上保護帯域内にある場合、前記バッファ充填度を減少させるか、同じに維持する前記符号化ビット・レートの変更だけを可能にし、前記バッファ充填度が前記下保護帯域内にある場合、前記バッファ充填度を増加させるか、同じに維持する前記符号化ビット・レートの変更だけを可能にするステップを含む、請求項17記載の方法。

【請求項19】前記符号化を動的に適応化するステップが、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御するス

テップを含み、前記符号化するステップが、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供するステップを含む、請求項1記載の方法。

【請求項20】前記符号化するステップが、前記多重ストリームの1ストリームを受信する複数のMPEG符号器を使用し、前記多重ストリームを並列に符号化するステップを含む、請求項1記載の方法。

【請求項21】前記多重ストリームを並列に符号化するステップが、多重圧縮ビデオ・ストリームを生成し、前記方法が前記多重圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネルを介して転送するステップを含む、請求項1記載の方法。

【請求項22】前記符号化を動的に適応化するステップが、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御するステップを含み、前記レート制御アルゴリズムが、前記符号化の結果生じる圧縮ビデオ・ストリームを受信するバッファの充填度に部分的にもとづき、前記少なくとも1つのストリームの前記符号化ビット・レートを制御する、請求項1記載の方法。

【請求項23】ビデオ・フレームの多重ストリームを処理する方法であって、各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを使用し、前記多重ストリームを並列に符号化するステップと、

前記符号化の間、前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出するステップと、

前記多重ストリームの各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記ストリームの前記符号化を動的に適応化するステップであって、前記動的に適応化するステップが、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整するステップを含み、前記符号化するステップが、前記多重ストリームから獲得される前記少なくとも1つの特性の相対変化に動的に適応する、適応化するステップとを含む、方法。

【請求項24】前記分析するステップが、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出するステップを含む、請求項23記載の方法。

【請求項25】前記符号化が多重圧縮ビデオ・ストリームを生成し、前記方法が前記各圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化するステップを含む、請求項24記載の方法。

【請求項26】前記多重化以前に前記各圧縮ビデオ・ストリームをバッファリングするステップを含み、前記分

析するステップが、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性に関する統計を導出するステップを含み、前記統計が使用されるビット、平均MQUANT、バッファ充填度、またはシーン変化に関する統計の少なくとも1つを含む、請求項25記載の方法。

【請求項27】前記符号化により使用される前記少なくとも1つの制御可能なパラメータが、前記ビデオ・フレームの各ストリームに対する符号化ビット・レートを含み、請求項23記載の方法。

【請求項28】前記符号化するステップが、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化するステップを含み、前記動的に適応化するステップが、前記ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットするステップを含む、請求項23記載の方法。

【請求項29】前記動的に適応化するステップが、前記ビデオ・フレームの各ストリームの前記GOPの各々に対して、符号化ビット・レートを決定するステップを含む、請求項28記載の方法。

【請求項30】前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たな前記GOPを開始するステップを含む、請求項29記載の方法。

【請求項31】前記少なくとも1つの特性が、前記ビデオ・フレームの各ストリームに対して、当該ストリームを構成するビデオ・データの複雑度を含み、前記分析するステップが、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測するステップを含む、請求項30記載の方法。

【請求項32】前記ビデオ・フレームの多重ストリームが、ビデオ・データの異なるソース、または多重ストリームに分割されるビデオ・データの共通ソースを含む、請求項23記載の方法。

【請求項33】ビデオ・フレームの多重ストリームを処理するシステムであって、

前記ビデオ・フレームの多重ストリームを並列に符号化し、前記多重ストリームの1ストリームを受信する複数の符号器と、

前記複数の符号器の各々に接続され、前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、少なくとも1つのストリームの符号化を動的に適応化する結合制御装置とを含む、システム。

【請求項34】前記結合制御装置が、シーン変化の検出に際して、前記少なくとも1つのストリームの符号化を動的に適応化する手段を含む、請求項33記載のシステム。

【請求項35】前記結合制御装置が、前記少なくとも1

10

20

30

40

50

つのストリームを符号化するために前記複数の符号器の1つにより使用される、少なくとも1つの制御可能な符号化パラメータを調整する手段を含む、請求項34記載のシステム。

【請求項36】前記制御可能な符号化パラメータの調整手段が、前記少なくとも1つのストリーム内のシーン変化の検出時に、または前記少なくとも1つのストリーム内の新たなフレーム・グループ(GOP)の開始時に、前記制御可能な符号化パラメータを調整する手段を含む、請求項35記載のシステム。

【請求項37】前記制御可能な符号化パラメータが符号化ビット・レートを含む、請求項36記載のシステム。

【請求項38】前記結合制御装置が、前記少なくとも1つのストリーム内のシーン変化の検出に際して、新たなフレーム・グループ(GOP)を開始する手段を含み、前記符号化ビット・レートが各GOP内で一定であり、前記少なくとも1つのストリームのGOP間で可変である、請求項33記載のシステム。

【請求項39】前記複数の符号器が多重圧縮ビデオ・ストリームを生成し、前記システムが、対応する前記符号器から出力を受信するように接続される複数のバッファを含み、前記結合制御装置が、前記少なくとも1つのストリームを符号化する前記符号器に接続される前記バッファの充填度に部分的にもとづき、前記少なくとも1つのストリームの符号化を動的に適応化する手段を含む、請求項33記載のシステム。

【請求項40】前記多重圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネル上に転送するマルチプレクサと、

前記マルチプレクサと前記一定ビット・レート・チャンネル間に接続され、前記一定ビット・レートでの多重化圧縮ビデオ・ストリームの転送を保証するチャンネル・バッファとを含む、請求項39記載のシステム。

【請求項41】前記結合制御装置が、前記チャンネル・バッファの充填度にもとづき、前記少なくとも1つのストリームの符号化を適応化する手段を含む、請求項40記載のシステム。

【請求項42】前記結合制御装置が、前記少なくとも1つのストリームの符号化ビット・レートを制御するレート制御アルゴリズムを含み、前記複数の符号器が、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供する、請求項33記載のシステム。

【請求項43】前記複数の符号器が、前記多重ストリームの1ストリームを受信するように並列に接続される複数のMPEG符号器を含み、前記多重ストリームがビデオ・データの異なるソース、または多重ストリームに分割されるビデオ・データの共通ソースを含む、請求項33記載のシステム。

【請求項44】ビデオ・フレームの多重ストリームを処

理するシステムであって、

少なくとも1つの制御可能な符号化パラメータを使用し、1つのストリームを符号化することにより、前記多重ストリームを並列に符号化する複数の符号器と、前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出する手段と、

前記多重ストリームの各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記各ストリームの前記符号化を動的に適応化する手段であって、前記動的に適応化する手段が、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整する手段を含み、前記符号化する手段が、前記多重ストリームから獲得される前記少なくとも1つの特性の相対変化に動的に適応する、適応化手段とを含む、システム。

【請求項45】前記分析手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出する手段を含む、請求項44記載のシステム。

【請求項46】前記複数の符号器が多重圧縮ビデオ・ストリームを生成し、前記システムが、前記圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネル上に転送するためのマルチプレクサを含む、請求項45記載のシステム。

【請求項47】前記複数の符号器の1つと前記マルチプレクサとの間に接続される複数のバッファと、

前記マルチプレクサと前記一定ビット・レート・チャンネル間に接続されるチャンネル・バッファとを含み、前記分析する手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性に関する統計を導出する手段を含み、前記統計が使用されるビット、平均MQANT、バッファ充填度、またはシーン変化に関する統計の少なくとも1つを含む、請求項46記載のシステム。

【請求項48】前記複数の符号器が、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化する手段を含み、前記動的に適応化する手段が、前記各ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットする手段を含む、請求項44記載のシステム。

【請求項49】前記動的に適応化する手段が、前記ビデオ・フレームの各ストリームの前記GOPの各々に対して、符号化ビット・レートを決定する手段を含む、請求項48記載のシステム。

【請求項50】前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たな前記GOPを開始する手段を含む、請求項4

9記載のシステム。

【請求項51】前記少なくとも1つの特性が、前記ビデオ・フレームの各ストリームに対して、当該ストリームを構成するビデオ・データの複雑度を含み、前記分析手段が、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測する手段を含む、請求項50記載のシステム。

【請求項52】ビデオ・フレームの多重ストリームの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、前記ビデオ・フレームの多重ストリームを符号化するように、前記コンピュータに指示する手段と、前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、少なくとも1つのストリームの符号化を動的に適応化するように、前記コンピュータに指示する手段とを含む、装置。

【請求項53】前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記少なくとも1つのストリーム内のシーン変化の検出時に、または前記少なくとも1つのストリーム内の新たなフレーム・グループ(GOP)の開始時に、前記少なくとも1つのストリームの符号化の動的な適応化を実行するように、前記コンピュータに指示する手段を含む、請求項52記載の装置。

【請求項54】前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記符号化から出力される圧縮ビデオ・ストリームを受信するように接続されるバッファの充填度に関するフィードバックを使用し、前記少なくとも1つのストリームの符号化を動的に適応化するように、前記コンピュータに指示する手段を含む、請求項53記載の装置。

【請求項55】前記符号化を動的に適応化するように前記コンピュータに指示する手段が、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御する手段を含み、前記符号化するように前記コンピュータに指示する手段が、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供するように、前記コンピュータに指示する手段を含む、請求項52記載の装置。

【請求項56】前記符号化するように前記コンピュータに指示する手段が、前記多重ストリームを並列に符号化し、多重圧縮ビデオ・ストリームを生成するように、前記コンピュータに指示する手段を含み、前記装置が前記多重圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化し、転送するように、前記コンピュータに指示する手段を含む、請求項52記載の装置。

【請求項57】ビデオ・フレームの多重ストリームの処

理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、ビデオ・フレームの各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを使用し、前記多重ストリームを並列に符号化するように、前記コンピュータに指示する手段と、

前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出するように、前記コンピュータに指示する手段と、各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記ストリームの前記符号化を動的に適応化するように、前記コンピュータに指示する手段であって、前記動的に適応化する手段が、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整する手段を含み、前記符号化手段が、前記多重ストリームから獲得される前記少なくとも1つの特性の相対変化に動的に適応する、適応化手段とを含む、装置。

【請求項58】前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出するように、前記コンピュータに指示する手段を含む、請求項57記載の装置。

【請求項59】前記符号化するように前記コンピュータに指示する手段が、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化するように、前記コンピュータに指示する手段を含み、前記動的に適応化する手段が、前記各ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットするように、前記コンピュータに指示する手段を含む、請求項57記載の装置。

【請求項60】前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たなフレーム・グループを開始するように前記コンピュータに指示する手段を含み、前記少なくとも1つの特性が、前記各ストリームのビデオ・フレームの複雑度を含み、前記分析するように前記コンピュータに指示する手段が、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測するように前記コンピュータに指示する手段を含む、請求項59記載の装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は一般に、デジタル・ビジュアル・イメージの圧縮に関して、特に、ビデオ・

データの多重ストリームを並列に動的に符号化し、一定ビット・レート・チャンネル上に多重化する制御技術に関する。この制御技術は、ビデオ・データの多重ストリームの相対複雑度に基づき、個々の符号化ビット・レートが各ビデオ・データ・ストリーム内で動的に調整されることを可能にする、単一バス技術を含む。

【0002】

【従来の技術】放送テレビジョン、通信ネットワーク、家電製品及びマルチメディア・コンピュータなどの業界における、様々なデジタル・ビデオ技術の出現は、急速に増えつつある。デジタル・ビデオ・アプリケーションのこの広範な利用は、デジタル情報の信号処理、編集及びデータ転送が、アナログ表現の処理に比較して、大変容易である事実により促進される。しかしながら、重要な点は、デジタル・ビデオはその普及を、近年デジタル・ビデオ圧縮のための作成された幾つかの規格に負っていることである。

【0003】デジタル・ビデオ圧縮ソリューションは、恐らく任意のデジタル・ビデオ・プラットフォームの最も重要な要素である。デジタル・ビデオは非圧縮形式では大量の情報を含むことがわかっているため、その操作、記憶及び伝送は、非常に時間を要し、高価となる。その結果、その内容の知覚品質を保存する一方、圧倒的なデータ量を低減するデジタル・ビデオ圧縮技術が考案された。互換のビデオ伸張方式が、データを再生のために圧縮解除するために使用される。

【0004】Moving Pictures and Expert Groupにより形成され、ISO/IEC 13818-2、"Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video 1996"で述べられるMPEG-2国際標準規格は、前述の業界のために、圧縮を標準化することを目的とする。ISOMPEG-2規格は、符号化ビットストリームの構文及び復号プロセスのセマンティクス（意味）を指定する。しかしながら、符号化パラメータの選択、及び性能対複雑度のトレードオフが、符号器開発者には残される。

【0005】MPEG-2符号化システムの効率性、固定通信帯域幅上で伝送される、またはデジタル媒体に記憶される知覚ビデオの精度により評価される。デジタル衛星システムなどの一部のアプリケーションでは、複数の番組が単一の大規模ストリームに多重化されて同報され、MPEG-2符号器のバンクが、全ての番組をモニタ及び符号化するために使用される一方で、受信チャンネルの品質を維持しようとする。MPEG-2ストリームは非同期転送モード（ATM）・ネットワークにおいて、固定通信帯域幅または動的帯域幅のいずれかにより送信される。

【0006】直接放送衛星（DBS）アプリケーションなどの典型的な統計的多重化システムでは、幾つかのビデオ・ビットストリーム（すなわち番組）が、単一定

ビット・レート・チャンネル上に多重化される。残念ながら、所定の一定ビット・レートでの各番組の符号化は、時間の経過に伴う番組のシーン内容の変化により、ピクチャ品質の劣化を生じ得る。ビデオ・ソースが異なる符号化を施されたり、異なるピクチャ・グループ（GOP）構造を有する場合など、符号器の作用が時間的に整列されない場合、問題は一層複雑になる。

【0007】

【発明が解決しようとする課題】使用可能な帯域幅を番組間で動的に分配することにより、システムの全体的なビデオ品質を最適化する多重番組圧縮技術が待望される。本発明は、符号化プロセスの間に、個々の符号器を動的に制御する結合レート制御技術を使用することにより、こうした技術を提供する。

【0008】

【課題を解決するための手段】要するに、1態様では、本発明の原理に従い、ビデオ・フレームの多重ストリームを処理する方法が提供される。この方法は、ビデオ・フレームの多重ストリームを並列に符号化するステップと、ビデオ・フレームの多重ストリームを構成するビデオ・フレームの相対複雑度に基づき、ビデオ・フレームの多重ストリームの少なくとも1つのストリームの符号化を、動的に適応化するステップとを含む。拡張された実施例では、符号化を動的に適応化するステップが、多重ストリームを構成するビデオ・フレームの複雑度に基づき、制御可能な符号化パラメータを調整するステップを含む。調整するステップは好適には、ビデオ・フレームの少なくとも1つのストリーム内のシーン変化の検出時に、または新たなピクチャ・グループの開始時に発生する。

【0009】別の態様では、ビデオ・フレームの多重ストリームを処理する方法が、ビデオ・フレームの各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを用いて、ビデオ・フレームの多重ストリームを符号化するステップと、符号化の間に、ビデオ・フレームの各ストリームを分析し、イントラフレーム特性またはインタフレーム特性の少なくとも1つに関する情報を導出するステップと、ビデオ・フレームの多重ストリームの各ストリームから獲得される少なくとも1つの特性に関する相対情報により、ビデオ・フレームの各ストリームの符号化を動的に適応化するステップとを含む。動的に適応化するステップは、ビデオ・フレームの各ストリームに対して、その符号化において使用される少なくとも1つの制御可能な符号化パラメータを調整するステップを含み、ビデオ・フレームの符号化が、ビデオ・フレームの多重ストリームから獲得される少なくとも1つの特性の相対変化に、動的に適応する。

【0010】更に別の態様では、ビデオ・フレームの多重ストリームを処理するシステムが提供される。このシステムは、ビデオ・フレームの多重ストリームを並列に

符号化する複数の符号器を含む。各符号器は、ビデオ・フレームの多重ストリームの1つのストリームを受信するように結合される。結合制御装置が各符号器に接続され、多重ストリームを構成するビデオ・フレームの相対複雑度にもとづき、ビデオ・フレームの多重ストリームの少なくとも1つのストリームの符号化を動的に適応化する。

【0011】更に別の態様では、本発明はビデオ・フレームの多重ストリームを処理するシステムを含む。このシステムは、ビデオ・フレームの多重ストリームを符号化する複数の符号器を含む。各符号器は、ビデオ・フレームの1つのストリームを符号化するための、少なくとも1つの制御可能な符号化パラメータを使用する。ビデオ・フレームの各ストリームを分析し、その少なくとも1つの特性に関する情報を導出する。この少なくとも1つの特性には、イントラフレーム特性またはインタフレーム特性の少なくとも1つが含まれる。ビデオ・フレームの各ストリームから獲得される少なくとも1つの特性に関する相対情報により、ビデオ・フレームの各ストリームの符号化を動的に適応化する手段が提供される。符号化を動的に適応化する手段は、ビデオ・フレームの各ストリームに対して、その符号化において使用される少なくとも1つの制御可能な符号化パラメータを調整する手段を含む。それによりビデオ・フレームの符号化が、ビデオ・フレームの多重ストリームから獲得される少なくとも1つの特性の相対変化に、動的に適応する。

【0012】更に別の態様では、ビデオ・フレームの多重ストリームの処理を指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む製造物が提供される。製造物内のコンピュータ読出し可能プログラム・コード手段が、コンピュータに、ビデオ・フレームの多重ストリームを符号化するように指示するコンピュータ読出し可能プログラム・コード手段と、コンピュータに、ビデオ・フレームの相対複雑度にもとづき、ビデオ・フレームの多重ストリームの少なくとも1つのストリームの符号化を動的に適応化するように指示するコンピュータ読出し可能プログラム・コード手段とを含む。

【0013】更に別の態様では、ビデオ・フレームの多重ストリームの処理を指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む製造物が提供される。製造物内のコンピュータ読出し可能プログラム・コード手段が、コンピュータに、ビデオ・フレームの各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを用い、ビデオ・フレームの多重ストリームを符号化するように指示するコンピュータ読出し可能プログラム・コード手段と、コンピュータにビデオ・フレームの各ストリームを分析し、イントラフレーム特性またはインタフレーム特性の少なくとも1つに関する情報を導出

するように指示する、コンピュータ読出し可能プログラム・コード手段と、コンピュータに、ビデオ・フレームの各ストリームから獲得された少なくとも1つの特性に関する相対情報を用い、ビデオ・フレームのストリームの符号化を動的に適応化するように指示する、コンピュータ読出し可能プログラム・コード手段とを含む。符号化の動的適応化は、ビデオ・フレームの各ストリームに対して、符号化において使用される少なくとも1つの制御可能な符号化パラメータを調整する。従って、ビデオ・フレームの各ストリームの符号化が、ビデオ・フレームの多重ストリームから獲得される少なくとも1つの特性の相対変化に、動的に適応する。

【0014】再度述べると、本発明は、例えばMPEG-2互換ビデオ符号器などを用い、幾つかのビデオ番組を並列に符号化する統計的多重化のためのシステム・ソリューションを提供する。符号器間でビット・レートを動的に割当てる結合レート制御技術が提供される。各符号器のビット・レートは、番組の相対複雑度、及び符号化される番組内で発生するシーン変化にもとづき決定される。この技術は、入力ビデオ信号の外部事前処理を必要としない。更に、ビデオ・ソースの符号化が、各符号器内において、同一のGOP構造及びGOP長を要求するように制限されない。有利な点として、各符号器はそのビット・レートをGOP境界において変更できる一方、GOP内では一定ビット・レートで作用する。全体として、この技術は区分的に一定であるが可変ビット・レートの圧縮をもたらす。符号器が異なるGOP長及び構造で作用でき、異なる時刻に符号化できる。従って、個々のビット・レートの総和が所定チャネル・ビット・レートよりも大きいかまたは小さい場合、時間間隔が存在し得る。更に、この技術は、シーン変化に対して提案され、システムのシーン変化に対する迅速な反応を保証することにより、ピクチャ品質を改善する。

【0015】

【発明の実施の形態】簡単に前述したように、本発明は複数のビデオ符号器を用い、ビデオ番組の多重ストリーム（例えばチャンネル）を並列に符号化するための統計的多重化制御技術に関する。符号器（MPEG準拠の符号器、及び前述のMPEG-2国際標準規格で述べられるような符号化プロセスを含み得る）の間で、ビット・レートを動的に割当てる外部結合レート制御技術が使用される。各符号器のビット・レートは、番組の相対複雑度にもとづき決定され、番組内のシーン変化及びGOP境界において調整が行われる。提案される技術は、入力ソースの外部事前処理を要求しない。図1及び図2は、本発明に従う制御技術において使用される単一ビデオ符号器の作用を示す。

【0016】説明のために、MPEG準拠符号化の一般的なフロー図を図1に示す。フロー図において、i番目のピクチャ及びi+1番目のピクチャのイメージが処理

され、動きベクトルを生成する。動きベクトルは、前の及び（または）後のピクチャ内で画素のマクロブロックが存在する場所を予測する。動きベクトルの使用は、MPEG規格における時間圧縮の主要な面である。図1に示されるように、一旦生成されると、動きベクトルは、i番目のピクチャからi+1番目のピクチャへの、画素のマクロブロック（MB）の変換のために使用される。

【0017】図1の符号化プロセスでは、i番目のピクチャ及びi+1番目のピクチャのイメージが符号器11内で処理され、動きベクトルが生成される。続くピクチャの入力イメージ111が、符号器11の動き推定ユニット43に入力する。動きベクトル113が動き推定ユニット43の出力として形成される。これらのベクトルは動き補償ユニット41により、先行及び（または）将来のピクチャから、“基準”データと呼ばれるマクロブロック・データを、このユニットの出力として取り出すために使用される。動き補償ユニット41の1出力が、動き推定ユニット43からの出力から減算され、離散コサイン変換器（DCT）21に入力される。離散コサイン変換器21の出力は、量子化器23内で量子化される。量子化器23の出力は2つの出力121及び131に分割され、一方の出力121はラン・レングス符号器などの下流の要素25に送られ、圧縮及び処理された後に伝送される。他の出力131は、画素の符号化マクロブロックの復元を通じて、フレーム・メモリ42に記憶される。図示の符号器では、この第2の出力131は逆量子化29及び逆離散コサイン変換31を通じて、差分マクロブロックの有損失バージョンを再構成する。このデータは動き補償ユニット41の出力と加算され、原ピクチャの有損失バージョンをフレーム・メモリ42に返却する。

【0018】図2に示されるように、3つのタイプのピクチャが存在する。“イントラ・ピクチャ”または“I”ピクチャは、独立に符号化されて伝送され、動きベクトルの定義を必要としない。これらの“I”ピクチャは動き推定のための基準イメージとして作用する。“予測ピクチャ”または“P”ピクチャは、先行ピクチャからの動きベクトルにより形成され、将来のピクチャの動き推定のための基準イメージとして作用する。最後に、“双方向ピクチャ”または“B”ピクチャは、2つの他のピクチャ、すなわち過去と将来のピクチャからの動きベクトルを用いて形成され、動き推定のための基準イメージとしては作用しない。

【0019】様々なピクチャ・タイプが符号器11により、幾つかのステップにおいて量子化される。第1に、ピクチャ・タイプに固有の量子化マトリックスが、8×8 DCTブロックに適用される。低周波数係数が重要視されるように、マトリックスの要素（重み）が選択される。これらの係数はより多くの情報を含み、ピクチャの知覚品質により大きく寄与する。次に、イメージの局所

的なアクティビティ、ピクチャ・タイプの複雑度、及びバッファ充填度にもとづき、量子化スケーリング・ファクタが計算される。この指標は、ピクチャに割当てられるビット数及びその知覚品質に直接関連付けられる。実際のビット数は、MPEG-2規格において定義されるプロシージャに従い、最初に各DCTブロックを走査し、次に量子化済み係数を一緒にグループ化することにより獲得される。各グループは、ハフマン・ルックアップ・テーブルを用いることにより、固有的可変長コード（VLC）によりエントロピ符号化される。差分パルス符号変調（DPCM）（動きベクトルの符号化に類似）及び一様量子化を用いて、DC係数が符号化される。

【0020】エントロピ符号器の出力は、エントロピ復号器に入力される。復号器の出力は、逆走査、逆量子化、及び逆離散コサイン変換を通じて、損失のある差分マクロブロックを再構成する。次に、復号されたピクチャが既知のように、遅延を通じて動き推定及び（または）動き補償に渡される。

【0021】MPEG-2符号器の動作機能については、Carrらによる1997年4月1日付けの米国特許出願第08/831157号、“Control Scheme For Shared-Use Dual-Port Predicted Error Array”で詳述されている。一定ビット・レート（CBR）符号化方式及び可変ビット・レート（VBR）符号化方式の基本が、N. Mo hsenianによる1998年3月19日付けの米国特許出願第09/044642号、“Real-Time Single Pass Variable Bit Rate Control Strategy And Encoder”で、ピクチャ・グループ（GOP）またはフレームに関連して詳述されている。更に、後述のようなイメージ統計を使用するフレーム・シーケンスの適応符号化が、Boiceらによる1998年3月20日付けの米国特許出願第09/046118号、“Adaptive Encoding Of A Sequence Of Still Frames Or Partially Still Frames Within Motion Video”で詳述されている。

【0022】前記の情報を背景として、本発明の原理に従う統計的多重化及びレート制御技術について、図3乃至図12を参照しながら説明することにする。

【0023】典型的な統計的多重化システムでは、幾つかのビデオ・ビットストリームが単一定ビット・レート・チャンネル上に多重化される。あいにく、所定の一定ビット・レートでの各番組の符号化は、時間に伴うシーン内容の変化により、ピクチャ品質の劣化を招き得る。従って、本発明に従う多重番組圧縮の根底の目的は、使用可能な帯域幅を番組間で動的に分配し、システムの全体的なビデオ品質を最大化することである。これは個々の符号器を制御する結合レート制御アルゴリズムを用い、全ての番組のピクチャ品質を等しくすることに対応する。本発明はビデオ・ソースの相対シーン内容に従い、符号器の間でビット・レートを割当てて一方、固定チャンネル・ビット・レートの要求を満足するものであ

る。

【0024】より詳細には、統計的多重化のためのシステム・ソリューションが、MPEG-2互換のビデオ符号器を使用し、幾つかのビデオ番組を並列に符号化するために提供される。符号器間で動的にビット・レートを割当てる外部結合レート制御アルゴリズムが述べられる。各符号器のビット・レートは、番組の相対複雑度及び番組内のシーン変化にもとづき決定される。本発明に従うシステム及び方法は、入力ソースのいかなる外部事前処理も必要としない。更に、ビデオ・ソースの符号化が各符号器内で、同一のGOP構造またはGOP長を有するように制限されない。各符号器は結合レート制御に従い、GOP境界において、そのビット・レートを変更する一方で、GOP内では一定ビット・レートで動作する。全体として、この技術は区分的に一定的可変ビット・レート圧縮をもたらす。実験結果によれば、本発明に従う多重番組ビデオ圧縮システムは、相対的に単純であるにも関わらず、外部事前処理無しに、良好なビクチャ品質をもたらす。更に、市販のMPEG-2符号器チップがこのシステムにおいて、成功裡に使用され得る。

【0025】基本的に、結合レート制御のために2つの異なるアプローチが可能であり、それらはフィードバック・アプローチと先読みアプローチである。フィードバック・アプローチでは、符号器により符号化プロセスの副産物として、統計が生成される。これらの統計が次に、続くビクチャに対するビット割当てを制御するために使用される。先読みアプローチでは、符号化の前に、統計がプリプロセッサにより計算され、次にこれらの統計がビクチャの符号化の前に、ビット・レートを調整するために使用される。いずれのアプローチでも、番組の複雑度を示す最適な統計を見い出すことが、課題となり得る。フィードバック・アプローチでは、統計は主として、符号化に関係する量に限られる。先読みアプローチはより大きな自由度を有するが、計算が複雑になったり、追加の装置が必要になったりする。

【0026】両方のアプローチにおいて、結合レート制御アルゴリズムが各ビクチャのビット・レートを計算し、個々の符号器の各々に対する全体レート制御を実行する。これは可変ビット・レート符号化をもたらす。通常、こうしたレート制御アルゴリズムは大抵、各符号器内で同一のGOP構造を要求するが、これは現実とかけ離れている。符号器は異なるGOP長及び構造において動作することができ、また動作する。

【0027】本発明によれば、フィードバック・アプローチにより符号器のビット・レートを動的に割当てる、統計的多重化のためのソリューションが開示される。レート制御技術が番組の相対複雑度にもとづき、符号器間でチャンネル・ビット・レートを分配する。番組の複雑度が、圧縮ビットストリームと共に、符号器により生成される符号化統計を用いて決定される。

【0028】結合レート制御アルゴリズムにより、各ビクチャのターゲット・ビット・レートを計算する代わりに、GOP境界において、またはシーン変化が発生する場合、ビット・レートが変更される。この技術は、符号器がGOP内において一定ビット・レートで動作することを可能にし、区分的に一定的可変ビット・レート圧縮をもたらす。符号器は同一のGOP構造を有する必要がなく、すなわち、GOP境界が各符号器内で異なる時刻に発生し得る。符号器の異なるGOP構造のため、チャンネル・バッファ及び対応するバッファ制御フィードバック・ループが、本発明に従いシステム内に組み込まれる。シーン変化に際して、GOP構造を動的に変更し、十分なビット・レート変化を可能にすることにより、迅速な反応が保証される。本システム及び方法は、例えばIBMの単一チップMPEG-2ビデオ符号器(部品番号: IBM39 MPEG S422 PBA 17C)を用いて実現され得る。

【0029】下記のセクション1は、本発明に従う多重番組ビデオ圧縮システムについて述べる。セクション2では、本発明の結合レート制御技術について、またセクション3では、シーン変化の場合の結合レート制御技術について述べる。セクション4では、チャンネル・バッファの最小サイズの決定、及び対応するチャンネル・バッファ制御について述べる。セクション5では、市販の単一チップMPEG-2符号器の、本システムにおける可能な使用について述べる。セクション6では、実験結果について示す。

【0030】1. 多重番組ビデオ圧縮システム: 図3は、本発明のフィードバック・アプローチを使用し、幾つかの番組(ソース1、ソース2、...、ソースn)を並列に符号化する多重番組ビデオ圧縮システム200を示す。システム200は、幾つかのMPEG-2ビデオ符号器210と、符号器210に接続される個々のバッファ220と、結合レート制御装置230と、マルチプレクサ240と、チャンネル・バッファ250とを含む。各符号器は、丁度符号化されたばかりのビクチャに関する統計を生成する。これらの統計は、結合レート制御アルゴリズム230への入力パラメータとなり、アルゴリズム230が後述のように、各個々の符号器のビット・レートを動的に計算する。

【0031】提案されるシステム200では、各符号器210のビット・レートが、番組(ソース1、ソース2、...、ソースn)の相対複雑度と、対応する番組内のシーン変化の発生にもとづき決定される。符号器は好適には、MPEG-2規格に互換のビットストリームを生成する。通常、符号器210の内部または外部のいずれかにおいて実行されるシーン変化検出以外は、入力ソースの追加の事前処理は要求されない。各符号器はそのビット・レートを結合レート制御装置230に従い、GOP境界またはシーン変化において変更する。こ

のことは符号器がGOP内では、MPEG-2規格に従うCBRビデオ・バッファ検証器モデルを使用し、一定ビット・レート(CBR)で動作することを可能にする。全体として、この技術は区分的に一定的可変ビット・レート圧縮をもたらす。

【0032】ビデオ・ソースの符号化が各符号器内で、同一のGOP構造またはGOP長を有するように制限されない。GOP境界は各符号器内で異なる時刻に発生し得、ビット・レート変化はGOP境界においてのみ有効となるので、チャンネル・バッファ250は、チャンネル・レート10の可能なオーバーフローまたはアンダーフローを補償するために使用される。チャンネル・バッファ制御フィードバック255もまた、結合レート制御アルゴリズム230に組み込まれ、チャンネル・バッファ250のオ*

$$R_i = R_o \cdot X_i / \left(\sum_i X_i \right)$$

1)

【0035】ここで R_i は番組 i のビット・レート、 R_o はチャンネル・レート、 X_i は番組 i の複雑度である。

【0036】ピクチャ複雑度の決定は、使用されるビット生成モデルにもとづき、これはMPEG-2テスト・*

$$b_i = c_i / Q_i$$

【0037】ここでモデル・パラメータ c_i は、ピクチャ内のターゲット・ビット数 b_i を生成するものであり、特定量子化スケール Q_i がセットされなければならない。式2)にもとづき、番組 i のビット・レートが、★

$$R_i = \frac{\sum_j (c_{ij} / Q_{ij})}{N_i / f_i}$$

3)

【0038】ここで c_{ij} はピクチャ j のビット生成モデル・パラメータであり、 Q_{ij} はピクチャ j の量子化パラメータであり、 N_i はGOP内のピクチャの数であり、 f_i は番組 i のフレーム・レートである。統計的多重化 ☆

$$\sum_i R_i \leq R_c$$

4)

【0039】全ての番組のピクチャ品質を等しくする目標を達成するために、理想的な量子化パラメータが式 ◆3) 及び式4) を用いて、次のように導出される。 ◆40 【数5】

$$Q_{ideal} = \frac{1}{R_c} \cdot \sum_i \left((f_i / N_i) \cdot \sum_j c_{ij} \right)$$

5)

【0040】この理想的な量子化パラメータは、各番組内の全てのピクチャに対して、等しいピクチャ品質をもたらすことができる。式3)において $Q_{i,j}$ を使用す *

$$R_i = R_c \cdot \left[(f_i / N_i) \cdot \sum_j c_{ij} \right] / \left[\sum_i \left((f_i / N_i) \cdot \sum_j c_{ij} \right) \right]$$

6)

【0041】本発明に従う統計的多重化システムにおいて、 c_{ij} は $b_{ij} \cdot Q_{ij}$ に等しく、 b_{ij} はピクチャを符号

* オーバーフローまたはアンダーフローを阻止する。

【0033】本発明の以下の説明では、結合レート制御技術、チャンネル・バッファの最小サイズの決定、及び対応するチャンネル・バッファ制御について詳述する。

【0034】2. 結合レート制御：ここで述べる結合レート制御技術はフィードバック概念にもとづき、そこでは符号器により生成された統計が、チャンネル・バッファ250の充填度に関する情報と共に、(圧縮ビットストリームと共に) 結合レート制御装置230に供給される。符号化される番組のビット・レートは、その番組の複雑度と、並列に符号化される全ての番組の複雑度の合計との比率に比例すると仮定される。すなわち、

【数1】

※モデル5で提案される次のモデルと類似である(ISO/IEC JTC1/SC29/WG11 N0400, "Test Model 5", April, 1993を参照のこと)。

【数2】

2)

★GOPを表示する時間間隔に対して次のように計算される。

【数3】

☆システムでは、その目標はチャンネル帯域幅を番組間で、次のように分配することである。

【数4】

化するために使用されるビットであり、 Q_{i1} は平均量子化パラメータである。更に、番組の複雑度が、GOPサイズのスライド式ウィンドウを用いて、ピクチャ複雑度の平均として推定される。

【0042】番組複雑度の関係により、たとえシーン変化が番組内において発生しなかった場合にも、ビット・レート変化がその番組内の任意のGOP境界において発生し得る。ビット・レート変化がシーン・カットの無い番組内で余りに突然発生する場合、たとえシステムの総品質が改善されても、ピクチャ品質はGOP間で多大に

変化し得る。この状況を阻止するために、ビット・レート変化がGOP境界上で制限されるが、シーン変化では、好適には制限は課せられない。例えば、シーン変化が発生しない場合、GOP境界における先行ビット・レートに対する10%の変化が許容され得る。シーン・カットが発生する場合、ビット・レート変化に対する制限は存在しない。

【0043】3. シーン変化における結合レート制御：ビデオ・データのストリーム内で、シーン変化はいつでも発生し得る。それらは任意のピクチャ・タイプにおいて、また任意のGOP位置において発生し得る。前述の議論では、ビット・レート変化がGOP境界においてだけ有効であり、符号器がGOP内で一定ビット・レート(CBR)・モードで作用することを仮定した。このアプローチの欠点は、番組の複雑度の変化に対するシステ

$$X_1 = \frac{X_I \cdot (1 + r_P \cdot n_P + r_B \cdot n_B)}{N_1 / f_1}$$

【0046】ここで X_1 はIピクチャの複雑度、 n_P 、 n_B はGOP内のPピクチャ及びBピクチャの数、 r_P 、 r_B はIピクチャ複雑度に対するそれぞれPピクチャ複雑度及びBピクチャ複雑度の比率である。 r_P 及び r_B の典型値は、それぞれ0.5及び0.25である。複雑度 X_1 は式6)でビット・レート計算のために使用される。新たなGOP内でより多くのピクチャが符号化されると、ピクチャを符号化するために使用される実際のビット及び平均量子化パラメータを適用することにより、複雑度が好適には連続的に更新される。GOPの符号化が継続すると、新たなシーンの複雑度の予測が、実際の符号化統計を使用することにより変化する。

【0047】既に、符号器がGOP内ではCBRモードで動作し、各符号器がCBRビデオ・バッファ検証器モデルを使用することについて述べた。バッファ・アンダーフローまたはオーバーフローは許可されない。例えば、CBRレート制御アルゴリズムの目標は、GOPの終りにおけるバッファ充填度が、初期バッファ充填度(例えばバッファ・サイズの80%)と同一であることを保証することである。しかしながら、これはターゲット※

$$R_{I\text{mod}} = R_1 + E \cdot f_1 / N_1$$

※ムの迅速な反応を阻止することである。この問題を克服するため、好適にはシーン変化状況において、特殊な処理が使用される。

【0044】符号器の内部または外部のいずれかにおいて、各番組の連続ピクチャ間のシーン変化検出分析が実行され、新たなシーン内の第1のピクチャを符号化する前に、シーン変化を知ることができる。シーン変化が発生する場合、現GOPが早まって終了される。1実施例では、新たなシーン内の第1のピクチャが、早まって終了されたGOPの最後のピクチャとして符号化され、その統計が新たなシーンの複雑度を予測するために、従って、式6)を用い、後続のGOPのビット・レートを計算するために使用される。図4は、原GOP構造の例、並びにシーン変化が発生する場合の、新たなGOP構造を示す。シーン変化が発生するピクチャ・タイプに応じて、3つのケースが区別されて示される。

【0045】新たなシーン複雑度の予測は、異なるピクチャ・タイプの経験的に決定された相対複雑度にもとづく。終りのGOPの最後のピクチャに当たる新たなシーンの第1のピクチャがPピクチャであり、あらゆるマクロブロックがイントラとして符号化される場合、このピクチャの複雑度はI複雑度と見なされる。このI複雑度にもとづき、新たなシーンの平均複雑度 X_1 が次のように推定される。

【数7】

7)

※ト・ビット・バジェットとIピクチャにつき使用される実際のビットとの不一致により、常に当てはまる訳ではない。GOP内のビットの過剰生成または過小生成により、バッファ充填度はそれぞれ初期レベルを下回ったり、上回ったりする。かなり大きなバッファ充填度誤差が累積し得る。この誤差は持ち越され、次のGOP内で補償されるように試行される。このレート制御は、GOP境界においてビット・レート変化がほとんど発生しないか、全く発生しない場合、良好に作用する。しかしながら、ビット・レート変化がシーン変化などにおいて突然発生する場合、本発明に従うバッファ充填度誤差(BFE)技術が、ピクチャ品質を改善するために適用される。

【0048】シーン変化が発生する場合、バッファ充填度誤差がシーン変化後の新たなGOPの開始において、0と見なされる。各符号器に接続されるバッファのアンダーフローまたはオーバーフローを阻止するために、式6)により計算されたビット・レートが次のように変更される。

【数8】

8)

【0049】ここで R_i は式6)に従い、番組 i に対して計算されたビット・レート、 E はバッファ充填度誤差ビット、 N_i はGOP内のピクチャの数、及び f_i は番組 i のフレーム・レートである。バッファ充填度誤差 E が正の場合(GOPの開始におけるバッファ充填度が初期値よりも小さい)、番組のビット・レートは増加され、 E が負の場合、番組のビット・レートは減少される。GOP内のレート制御のために、初期バッファ充填度は、例えばバッファ・サイズの80%を使用する代わりに、先行GOPの終りにおけるバッファ充填度と見なされる。BFE技術は、シーン変化後のピクチャ品質を改善する。

【0050】4. チャンネル・バッファ・サイズ及びフィードバック制御：符号器が異なるGOP長及び構造にて動作可能であること、または異なる時刻に符号化を開始*

$$B_s = 2 \cdot \Delta R_{\max} \cdot t_{\text{gop}_{\max}}$$

【0052】ここで $\Delta R_{\max} = \Sigma R_i - R_c$ 、 $t_{\text{gop}_{\max}}$ は最大GOP時間である。

【0053】式9)では、係数2が使用されている。なぜなら、チャンネル・ビット・レートの過小生成及び過剰生成が仮定され、バッファが最初にこのサイズ B_s の半分まで充填され、その後、 R_c のレートでビットを連続的に出力するからである。この場合、初期遅延は、バッファをそのサイズの半分まで充填するために要求される時間に対応する。例えば、チャンネル・バッファ・レートが16Mビット/秒、 ΔR_{\max} が8Mビット/秒、及び $t_{\text{gop}_{\max}} = 0.5$ 秒の場合、30フレーム/秒のフレーム・レートでの最小バッファ・サイズは8Mビットで、対応する初期遅延は0.25秒である。統計的多重*

$$0 < B_r(t) < B_s$$

【0055】パラメータ' a 'が保護帯域のサイズを決定し、例えば0.25として選択される。

【0056】実際のバッファ充填度 B_r に応じて、区別可能な3つのケースが存在する。これらのケースについて、対応するビット・レートの変更と共に、次に示す★

$$a \cdot B_s \leq B_r \leq (1-a) \cdot B_s$$

【0058】この場合では、番組のために計算されたビット・レートが、極端な状況を除き、一般に変更されない。

★

$$R_i = R_i \cdot [R_c + ((1-a) B_s - B_r) / t_{\text{gop}_{\max}}] / (\Sigma R_i)$$

(オーバーフロー無し)

12)

【数13】 $\Sigma R_i < R_c$ 且つ $R_c - \Sigma R_i > B_r / t_{\text{gop}_{\max}}$ の場合、

$$R_i = R_i \cdot [R_c - (B_r - a B_s) / t_{\text{gop}_{\max}}] / (\Sigma R_i)$$

(アンダーフロー無し)

13)

【0059】前記以外では、何もアクションは実行されない。

【0060】ケース2：これはバッファ充填度が上側の*

$$B_r > (1-a) \cdot B_s$$

14)

【0061】この場合では、バッファ充填度を減少させるか、または B_r に留まるビット・レート変化だけが許

＊し得る事実により、個々のビット・レートの総和が所定チャンネル・ビット・レートよりも大きかったり、小さかったりする場合、時間間隔が存在し得る。これを改善するため、チャンネル・バッファが要求され、正確にチャンネル・ビット・レートでビットストリームを出力する。このバッファに関連して、2つの問題が考慮されなければならない。それらはバッファの最小サイズの決定と、チャンネル・バッファ・アンダーフロー及びオーバーフローを阻止するための制御技術である。

10 【0051】チャンネル・ビット・レートからの最大合計偏差が ΔR_{\max} であり、最悪の場合として、この偏差の最大期間が、符号器間の最長GOP時間と同じであると仮定する。この場合、チャンネル・バッファの要求最小サイズは次のように決定される。

【数9】

9)

＊化システムにおいて、式9)により決定されるよりも小さなチャンネル・バッファが所望される場合、チャンネル・ビット・レートからの最大合計偏差は、それに従い制限されなければならない。

20 【0054】チャンネル・バッファ・アンダーフローまたはオーバーフローを阻止するために、バッファ・モデル(図5参照)が好適には使用される。チャンネル・バッファ・モデルは、バッファの頂部及び底部に所定の保護帯域を含む。これらの保護帯域はビット・レートの分配を調整するために使用される。アンダーフロー及びオーバーフローを阻止するために、任意の時刻におけるバッファ充填度 B_r が次の条件を満足しなければならない。

【数10】

10)

★とにする。

【0057】ケース1：これはバッファ充填度が保護帯域間にある場合である。すなわち、

【数11】

11)

☆【数12】 $\Sigma R_i > R_c$ 且つ $\Sigma R_i - R_c > (B_s - B_r) / t_{\text{gop}_{\max}}$ の場合、

＊保護帯域内にある場合である。すなわち、

【数14】

可される。

50 【数15】 $\Sigma R_i > R_c$ の場合、

23

$$R_i = R_i \cdot (R_c / (\sum R_i))$$

(スケーリング・ダウン)

15)

【数16】 $\sum R_i < R_c$ 且つ $R_c - \sum R_i > B_r / \text{t g o p} * * \dots$ の場合、

$$R_i = R_i \cdot [R_c - (B_r - a B_r) / \text{t g o p} \dots] / (\sum R_i)$$

(アンダーフロー無し)

16)

【0062】前記以外では、何もアクションは実行されない。 ※保護帯域内にある場合である。すなわち、

【数17】

【0063】ケース3：これはバッファ充填度が下側の※

$$B_r < a \cdot B_r$$

17)

【0064】この場合では、バッファ充填度を増加させ 10★可される。

るか、または B_r に留まるビット・レート変化だけが許 ★ 【数18】 $\sum R_i < R_c$ の場合、

$$R_i = R_i \cdot (R_c / (\sum R_i))$$

(スケーリング・アップ)

18)

【数19】 $\sum R_i > R_c$ 且つ $\sum R_i - R_c > (B_r - B_r) / \star \star \text{t g o p} \dots$ の場合、

$$R_i = R_i \cdot [R_c + ((1-a) B_r - B_r) / \text{t g o p} \dots] / (\sum R_i)$$

(オーバーフロー無し)

19)

【0065】前記以外では、何もアクションは実行されない。

【0066】5. 統計的多重化システム内でのMPEG-2符号器の使用：本発明の原理に従う多重番組ビデオ圧縮システム内で使用される全てのMPEG-2符号器は、少なくとも、前述の結合レート制御アルゴリズムにより要求される必要な統計を提供できなければならない。符号器はまた、GOP境界においてビット・レートを変更する能力を有さねばならない。本発明の利点を更に利用するために、シーン変化に際して、符号器はGOP構造を動的に変更し、バッファ充填度誤差を計算するための量を提供し、シーン変化検出及び反応を内部的または外部的に実行できなければならない。

【0067】6. 実験結果：本発明に従い構成されるシステムの性能を立証するために、異なるシーン内容のイメージ・シーケンスを使用することにより、幾つかの実験がシミュレーションを通じて実行された。開発された多重番組ビデオ圧縮システムは、4つのMPEG-2符号器(符号器1乃至符号器4)を用いてシミュレートされた。符号器の各々は、要求されるイメージ統計を出力する能力を有した。シーン変化検出は、符号器内で実行された。ビデオ・ソースは、それらの各々が、例えばスポーツ・シーン、自然、頭と肩だけのシーンなどの異なるシーン内容を表し、各々がシーン変化を含むように選択された。

【0068】第1のビデオ・ソースのセットが、4：2：2クロマ形式の日本IBMコマーシャル、卓球、花園及び移動カレンダー(移動カレンダーを伴う花園)、並びにMTVロゴを有する保護(Care)であった。ソースは4：2：2クロマ形式で符号化された。2つのBピクチャがアンカ・ピクチャ間に配置された。閉GOP長は、符号器1及び2では16として、符号器3及び4では13として選択された。フレーム・レートは各符号器において、29.97フレーム/秒であった。チャンネル・レ

ートは16Mビット/秒で、チャンネル・バッファ・サイズは式9)に従い、8Mビットであった。各符号器は4Mビット/秒のビット・レートで符号化を開始した。この初期ビット・レートは、本発明の結合レート制御技術に従い、動的に変更された。

【0069】図6は、結合レート制御装置を用い、各符号器に動的に割当てられるビット・レートを示す。日本IBMコマーシャル及びMTVシーケンスが、他の2つのソースに対して低いビット・レートを有することがわかる。

【0070】第1のビデオ・ソースのセットを用い、4つの符号器により生成された合計ビット・レートが図7に示される。このグラフは、16Mビット/秒におけるチャンネル・ビット・レートの過小生成または過剰生成を示し、チャンネル・バッファ及びフィードバック制御の必要性を立証している。

【0071】本提案のシステムの性能が、各符号器がそのソースを固定ビット・レートで符号化する方式と比較された。これらの符号器はシーン変化検出をそれら自身実行する。表1は、第1のビデオ・シーケンスのセットに対して、本発明により達成される平均PSNR値と、CBRモードにおいて、固定ビット・レート(4Mビット/秒)により達成される平均PSNR値とを示す。表1が示すように、簡単なソース(日本IBM、MTV)は、統計的多重化システムにおいては、固定4Mビット/秒レートよりも僅かに低い品質で符号化された。しかしながら、これは本提案のシステムにおいて、より複雑なソース(卓球、花園及び移動カレンダー)を、固定CBRモードの場合よりも高い品質で符号化することを可能にする。従って、本発明に従う統計的多重化システムは、固定ビット・レートの場合よりも優れた全体ピクチャ品質を提供した。

【表1】

ソース	平均PSNR [dB]	
	統計的多重化 ($R_c = 16\text{Mビット/秒}$)	CBR (4Mビット/秒)
日本IBM (符号器1)	38.48	40.11
卓球 (符号器2)	32.11	31.29
花国及び移動カレンダー (符号器3)	30.26	28.24
MTV (符号器4)	37.79	38.65

【0072】表2は、同一のビデオ・ソースを使用したとき、チャンネル・ビット・レートが32Mビット/秒で、16Mビットのチャンネル・バッファを有する場合、*及び8Mビット/秒のCBRモードにおける固定ビット・レートの場合の平均PSNR値を示す。【表2】

ソース	平均PSNR [dB]	
	統計的多重化 ($R_c = 32\text{Mビット/秒}$)	CBR (8Mビット/秒)
日本IBM (符号器1)	40.49	42.40
卓球 (符号器2)	35.36	34.61
花国及び移動カレンダー (符号器3)	34.16	31.70
MTV (符号器4)	39.96	41.00

【0073】チャンネル・バッファ・モード及びフィードバック制御の有効性を立証するために、図8は、32Mビット/秒のチャンネル・ビット・レートにおけるシーケンスの符号化の間の、チャンネル・バッファ充填度を示す。図示のように、符号化の間に、バッファ・アンダーフローまたはオーバーフローは発生しなかった。

【0074】表3は、16Mビット/秒のチャンネル・ビット・レートにおいて、本提案のバッファ充填度誤差 ※

※(BFE)技術を使用する場合の、及び使用しない場合のシーン変化後の第1のピクチャに対するPSNR値を含む。表3が示すように、ここで述べられるBFE技術を使用することにより、それを使用しないアルゴリズムに比較して、約0.64dB乃至約2.17dBのPSNRの改善が達成された。

【表3】

	PSNR [dB]				PSNR [dB]			
	符号器1				符号器3			
ピクチャ	I_{11}	B_{11}	B_{12}	P_{11}	I_{13}	B_{13}	B_{14}	P_{13}
BFE技術有	34.79	36.75	36.44	35.80	39.86	40.07	40.37	40.18
BFE技術無	34.15	36.09	35.77	35.09	38.72	39.08	39.30	38.99
	符号器2				符号器4			
	I_{12}	B_{12}	B_{13}	P_{12}	I_{14}	B_{14}	B_{15}	P_{14}
BFE技術有	30.49	33.52	33.49	32.55	39.00	39.09	38.48	38.77
BFE技術無	28.55	31.35	31.50	30.38	38.03	38.20	37.45	37.60

【0075】第2の実験のセットでは、IBMワールドブック・コマーシャル(符号器1)、Mixd(符号器2)、フットボール(符号器3)、及びMixe(符号器4)が、入力ソースとして使用された。Mixdはバイク、摩天楼及びバスケットボールのシーケンスを含

み、Mixeでは、スージ(Susie)・シーケンスに、森の中の小屋のシーンが続く。これらのソースは第1のビデオ・ソースのセットよりも、幾分複雑であるので、チャンネル・ビット・レートは24Mビット/秒に選択され、チャンネル・バッファは12Mビットであった。符号

化パラメータは、閉GOP長が符号器1及び符号器2では13で、符号器3及び符号器4では16であった以外、第1の実験のセットの場合と同一であった。CBRケースでは、ビット・レートが6Mビット/秒に固定された。図9は、本発明の結合レート制御技術に従う、符号器の動的ビット・レート変化を示す。第2のビデオ・ソースのセットを符号化するための合計ビット・レートが図10に示される。これは2.4Mビット/秒におけるチャンネル・ビット・レートの過小生成及び過剰生成を示す。

*10 【表4】

ソース	平均PSNR [dB]	
	統計的多重化 ($R_c = 24\text{M}$ ビット/秒)	CBR (6Mビット/秒)
IBMワールドブック (符号器1)	37.03	37.72
Mixed (符号器2)	33.74	33.19
フットボール (符号器3)	37.67	37.58
Mixed (符号器4)	38.68	39.37

【0077】2.4Mビット/秒のチャンネル・ビット・レートにて、この第2のビデオ・ソースのセットを符号化するためのチャンネル・バッファ充填度が、図11に示される。バッファ・アンダーフローまたはオーバーフローは存在しない。

【0078】7. プロセス概要：図12は、本発明に従う符号化処理の概要を示す。統計が個々の符号器からフィードバックとして、レート制御装置に供給される。これらの統計は使用されるビット、平均MQUANT、及びバッファ充填度を含み得る（ステップ300）。更に、ビデオ・フレームの各ストリームにおいて発生するシーン変化、チャンネル・バッファ充填度、及び前述のようにシステム内において発生する最大GOP長に関する情報が提供される。チャンネル・バッファ充填度は、チャンネル・バッファ自身からフィードバック信号として受信される。ビット・レート制御装置は初期に、シーン変化を示すシーン変化フラグがセットされたか否かを判断する（ステップ310）。セットされた場合、新たなシーンの最初のフレームが、前のシーンの最後のフレームとして、または新たな1ピクチャGOPの最初のピクチャとして符号化される（ステップ320）。新たなシーンのこの最初のピクチャは、次に新たなシーンの複雑度を予測するために使用される（ステップ330）。

【0079】ビット・レート R_i が各符号器に対して計算される（ステップ340）。次に、シーン変化フラグがセットされたか否かを判断し（ステップ350）、セットされた場合、バッファ充填度誤差技術により、符号器のビット・レート R_i を変更する（ステップ360）。次に、チャンネル・バッファ制御技術が前述のように実行され（ステップ370）、現フレームが、符号器 i にて符号化されるビデオ・フレームのストリームのG

*【0076】表4は、本提案システムにより、及びCBRモードにおける固定ビット・レート（6Mビット/秒）により達成された、平均PSNR値を含む。この表は、第1の実験のセットに対して達成されたビジュアル品質と、同一の傾向を示す。本発明の統計的多重化システムは、CBRモードと比較して、単純なソース（IBMワールドブック、Mixed）に対しては、僅かに低いビクチャ品質をもたらしたが、より複雑なソース（Mixed、フットボール）の品質を改善した。

20 OP境界にあるか否かを判断する（ステップ380）。否定の場合、処理はステップ300に戻り、入力パラメータの次のセットを収集する。現フレームがGOP境界にある場合、新たなビット・レート R_i が符号器 i に伝達される（ステップ390）。

【0080】再度要約すると、本発明によれば、複数の符号器間でビット・レートを動的に割当て、ビデオ・ストリームを並列に符号化し、共通チャンネル上に多重化する結合レート制御技術が提供される。番組の相対複雑度及び番組内で発生するシーン変化にもとづき、チャンネル・ビット・レートが符号器間で分配される。ここで述べられる方法は、入力ソースの外部事前処理を要求しない。番組の複雑度が符号化統計及びシーン変化検出にもとづき決定され、それらが符号器により圧縮ビットストリームと一緒に生成される。各ビクチャのターゲット・ビット・レートを計算する代わりに、ビット・レートがGOP境界において、またはシーン変化が発生する場合、変更される。これは符号器がGOP内では、一定ビット・レート（CBR）で動作することを可能にする。

【0081】GOP境界でのみビット・レートを変更することは、番組の複雑度変化に対する結合レート制御装置の反応時間を制限する。この問題を解決するために、シーン変化検出が、ビデオ・フレームのストリーム内の各連続ピクチャ間で実行されると仮定し、またそれが新たなシーン内の最初のフレームの符号化以前に知れていると仮定する。更に、符号器がGOP構造を動的に変更可能であると仮定する。これは市販の符号器において使用可能な機能である。シーン変化が発生する場合、現GOPが早まって終了され、新たなシーンの最初のピクチャが符号化される。新たなシーンのこの最初のピクチャからの統計が、次に新たなシーンの複雑度を予測するた

めに使用され、結果的に、続くGOPのビット・レート
を計算する。

【0082】シーン変化の場合にピクチャ品質を改善
するために、更に追加の変更が開発された。特に、GOP
内のビットの過剰生成または過小生成に起因するバッ
ファ充填度誤差が、シーン変化後の新たなGOPの開始に
おいて、0と見なされる。各符号器に接続されるバッ
ファ内のアンダーフローまたはオーバーフローを阻止す
るために、結合レート制御装置により計算されるビット
・レートが変更される。バッファ充填度誤差が正の場合
(GOPの開始におけるバッファ充填度が、初期バッ
ファ充填度(例えばバッファ・サイズの80%)よりも小
さい)、番組のビット・レートが増加される。反対に、
バッファ充填度誤差が負の場合、番組のビット・レート
が減少される。この技術はシーン変化の後、より優れた
ピクチャ品質をもたらす。この概念は更に、動的GOP
構造機能を有さない符号器、及びシーン変化検出が外部
的に実行される符号器に対して拡張され得る。

【0083】符号器は好適には同一のフレーム・レート
で動作するが、異なるGOP長及び構造を有し得、異な
る時刻に符号化を開始し得る。従って、個々のビット
・レートの総和が、所定のチャンネル・ビット・レートより
も大きかったり、小さかったりする場合、時間間隔が存
在し得る。チャンネル・バッファがシステム内に組み込ま
れ、正確にチャンネル・ビット・レートで多重化ビット
・ストリームを出力する。また、チャンネルの最小サイズを
決定する方法も開発された。これはチャンネル・ビット
・レートからの許容合計偏差、及びシステム内で発生する
最大GOP時間にもとづく。チャンネル・バッファ・アン
ダーフローまたはオーバーフローを阻止するために、バ
ッファ・フィードバックが結合レート制御装置内に組み
込まれる。

【0084】本発明は例えば、コンピュータ使用可能媒
体を有する製造物(例えば1つ以上のコンピュータ・プ
ログラム製品)内に含まれ得る。この媒体は、例えば本
発明の機能を提供し、容易にするコンピュータ読出し可
能プログラム・コード手段を実現する。製造物はコンピ
ュータ・システムの一部として含まれるか、或いは別々
に販売され得る。

【0085】更に、本発明の機能を実行するために、マ
シンにより実行可能な命令の少なくとも1つのプログラ
ムを実現する、マシンにより読出し可能な少なくとも1
つのプログラム記憶装置が提供され得る。

【0086】まとめとして、本発明の構成に関して以下
の事項を開示する。

【0087】(1)ビデオ・フレームの多重ストリーム
を処理する方法であって、前記ビデオ・フレームの多重
ストリームを並列に符号化するステップと、前記多重ス
トリームを構成する前記ビデオ・フレームの相対複雑度
にもとづき、少なくとも1つのストリームの符号化を動

的に適応化するステップとを含む、方法。

(2)前記少なくとも1つのストリームの符号化を動的
に適応化するステップを、当該ストリーム内のシーン変
化の検出に際して実行するステップを含む、前記(1)
記載の方法。

(3)前記符号化を動的に適応化するステップが、前記
少なくとも1つのストリーム内の前記シーン変化の検出
に際して、当該ストリームを符号化するために使用され
る制御可能な符号化パラメータを調整するステップを含
む、前記(2)記載の方法。

(4)前記少なくとも1つのストリーム内の前記シーン
変化を検出するステップを含む、前記(3)記載の方法。

(5)前記符号化を動的に適応化するステップが、前記
多重ストリームを構成する前記ビデオ・フレームの相対
複雑度にもとづき、前記多重ストリームの各ストリーム
の符号化を動的に適応化するステップを含む、前記
(1)記載の方法。

(6)前記符号化を動的に適応化するステップが、前記
符号化により使用される制御可能な符号化パラメータを
調整し、前記各ストリームを符号化するステップを含
む、前記(5)記載の方法。

(7)前記各ストリームに対して、前記制御可能な符号
化パラメータを調整する前記ステップが、当該ストリー
ム内のシーン変化の検出時に、または新たなフレーム・
グループ(GOP)の開始時に発生する、前記(6)記
載の方法。

(8)前記ストリーム内のシーン変化の検出無しに、新
たなフレーム・グループが開始するとき、前記制御可能
な符号化パラメータの調整を、所定の割合調整に制限す
るステップを含む、前記(7)記載の方法。

(9)シーン変化の検出に際して、前記符号化を動的に
適応化するステップが、現フレーム・グループ(GO
P)を終了し、新たなGOPを開始するステップを含
み、前記制御可能な符号化パラメータを調整するステッ
プが、前記所定の割合調整を使用することなく、前記制
御可能な符号化パラメータを調整するステップを含む、
前記(8)記載の方法。

(10)前記多重ストリームの隣接ビデオ・フレームを
モニタすることにより、前記シーン変化を検出するステ
ップを含む、前記(9)記載の方法。

(11)前記各ストリームに対する前記制御可能な符号
化パラメータが、符号化ビット・レートを含む、前記
(7)記載の方法。

(12)前記符号化が前記各ストリームに対して、当該
ストリーム内のシーン変化の検出時に、新たなフレーム
・グループ(GOP)を開始するステップを含み、前記
各ストリームに対して、前記符号化ビット・レートが各
GOP内で一定であり、前記GOP間で可変である、前
記(11)記載の方法。

(13) 前記符号化が多重圧縮ビデオ・ストリームを生成し、前記方法が、各圧縮ビデオ・ストリームをバッファ内にバッファリングするステップを含み、前記符号化を動的に適応化するステップが、前記各ストリームに対して、前記圧縮ビデオ・ストリームを受信する前記バッファの充填度にもとづき、前記符号化ビット・レートを変更するステップを含む、前記(11)記載の方法。

(14) 前記符号化ビット・レートを変更するステップが、前記符号化ビット・レートを

$$R_{i,old} = R_i + E \cdot f_i / N_i$$

として変更するステップを含み、ここで R_i はビデオ・フレームのストリーム i に対して計算されたビット・レート、 E はバッファ充填度誤差ビット、 N_i はストリーム i のGOP内のピクチャ数、 f_i はストリーム i のフレーム・レートである、前記(13)記載の方法。

(15) 前記多重圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化するステップを含む、前記(13)記載の方法。

(16) 前記多重圧縮ビデオ・ストリームを、前記一定ビット・レート・チャンネルに接続されるチャンネル・バッファ内にバッファリングするステップを含み、前記符号化を動的に適応化するステップが、前記各ストリームに対して、前記チャンネル・バッファの充填度にもとづき、前記符号化ビット・レートを変更するステップを含む、前記(15)記載の方法。

(17) 前記チャンネル・バッファの上限及び下限において、保護帯域を事前に定義するステップを含み、前記バッファ充填度が前記チャンネル・バッファの前記上下のいずれかの保護帯域内にある場合、前記符号化ビット・レートを変更するステップを含む、前記(16)記載の方法。

(18) 前記符号化ビット・レートを変更するステップが、前記バッファ充填度が前記上保護帯域内にある場合、前記バッファ充填度を減少させるか、同じに維持する前記符号化ビット・レートの変更だけを可能にし、前記バッファ充填度が前記下保護帯域内にある場合、前記バッファ充填度を増加させるか、同じに維持する前記符号化ビット・レートの変更だけを可能にするステップを含む、前記(17)記載の方法。

(19) 前記符号化を動的に適応化するステップが、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御するステップを含み、前記符号化するステップが、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供するステップを含む、前記(1)記載の方法。

(20) 前記符号化するステップが、前記多重ストリームの1ストリームを受信する複数のMPEG符号器を使用し、前記多重ストリームを並列に符号化するステップを含む、前記(1)記載の方法。

(21) 前記多重ストリームを並列に符号化するステップが、多重圧縮ビデオ・ストリームを生成し、前記方法が前記多重圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネルを介して転送するステップを含む、前記(1)記載の方法。

(22) 前記符号化を動的に適応化するステップが、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御するステップを含み、前記レート制御アルゴリズムが、前記符号化の結果生じる圧縮ビデオ・ストリームを受信するバッファの充填度に部分的にもとづき、前記少なくとも1つのストリームの前記符号化ビット・レートを制御する、前記(1)記載の方法。

(23) ビデオ・フレームの多重ストリームを処理する方法であって、各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを使用し、前記多重ストリームを並列に符号化するステップと、前記符号化の間、前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出するステップと、前記多重ストリームの各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記ストリームの前記符号化を動的に適応化するステップであって、前記動的に適応化するステップが、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整するステップを含み、前記符号化するステップが、前記多重ストリームから獲得される前記少なくとも1つの特性の相対変化に動的に適応する、適応化するステップとを含む、方法。

(24) 前記分析するステップが、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出するステップを含む、前記(23)記載の方法。

(25) 前記符号化が多重圧縮ビデオ・ストリームを生成し、前記方法が前記各圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化するステップを含む、前記(24)記載の方法。

(26) 前記多重化以前に前記各圧縮ビデオ・ストリームをバッファリングするステップを含み、前記分析するステップが、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性に関する統計を導出するステップを含み、前記統計が使用されるビット、平均MQUANT、バッファ充填度、またはシーン変化に関する統計の少なくとも1つを含む、前記(25)記載の方法。

(27) 前記符号化により使用される前記少なくとも1つの制御可能なパラメータが、前記ビデオ・フレームの各ストリームに対する符号化ビット・レートを含む、前記(23)記載の方法。

(28) 前記符号化するステップが、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化するステップを含み、前記動的に適応化するステップが、前記ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットするステップを含む、前記(23)記載の方法。

(29) 前記動的に適応化するステップが、前記ビデオ・フレームの各ストリームの前記GOPの各々に対して、符号化ビット・レートを決定するステップを含む、前記(28)記載の方法。

(30) 前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たな前記GOPを開始するステップを含む、前記(29)記載の方法。

(31) 前記少なくとも1つの特性が、前記ビデオ・フレームの各ストリームに対して、当該ストリームを構成するビデオ・データの複雑度を含み、前記分析するステップが、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測するステップを含む、前記(30)記載の方法。

(32) 前記ビデオ・フレームの多重ストリームが、ビデオ・データの異なるソース、または多重ストリームに分割されるビデオ・データの共通ソースを含む、前記(23)記載の方法。

(33) ビデオ・フレームの多重ストリームを処理するシステムであって、前記ビデオ・フレームの多重ストリームを並列に符号化し、前記多重ストリームの1ストリームを受信する複数の符号器と、前記複数の符号器の各々に接続され、前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、少なくとも1つのストリームの符号化を動的に適応化する結合制御装置とを含む、システム。

(34) 前記結合制御装置が、シーン変化の検出に際して、前記少なくとも1つのストリームの符号化を動的に適応化する手段を含む、前記(33)記載のシステム。

(35) 前記結合制御装置が、前記少なくとも1つのストリームを符号化するために前記複数の符号器の1つにより使用される、少なくとも1つの制御可能な符号化パラメータを調整する手段を含む、前記(34)記載のシステム。

(36) 前記制御可能な符号化パラメータの調整手段が、前記少なくとも1つのストリーム内のシーン変化の検出時に、または前記少なくとも1つのストリーム内の新たなフレーム・グループ(GOP)の開始時に、前記制御可能な符号化パラメータを調整する手段を含む、前記(35)記載のシステム。

(37) 前記制御可能な符号化パラメータが符号化ビット・レートを、前記(36)記載のシステム。

(38) 前記結合制御装置が、前記少なくとも1つのス

トリーム内のシーン変化の検出に際して、新たなフレーム・グループ(GOP)を開始する手段を含み、前記符号化ビット・レートが各GOP内で一定であり、前記少なくとも1つのストリームのGOP間で可変である、前記(33)記載のシステム。

(39) 前記複数の符号器が多重圧縮ビデオ・ストリームを生成し、前記システムが、対応する前記符号器から出力を受信するように接続される複数のバッファを含み、前記結合制御装置が、前記少なくとも1つのストリームを符号化する前記符号器に接続される前記バッファの充填度に部分的にもとづき、前記少なくとも1つのストリームの符号化を動的に適応化する手段を含む、前記(33)記載のシステム。

(40) 前記多重圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネル上に転送するマルチプレクサと、前記マルチプレクサと前記一定ビット・レート・チャンネル間に接続され、前記一定ビット・レートでの多重化圧縮ビデオ・ストリームの転送を保証するチャンネル・バッファとを含む、前記(39)記載のシステム。

(41) 前記結合制御装置が、前記チャンネル・バッファの充填度にもとづき、前記少なくとも1つのストリームの符号化を適応化する手段を含む、前記(40)記載のシステム。

(42) 前記結合制御装置が、前記少なくとも1つのストリームの符号化ビット・レートを制御するレート制御アルゴリズムを含み、前記複数の符号器が、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供する、前記(33)記載のシステム。

(43) 前記複数の符号器が、前記多重ストリームの1ストリームを受信するように並列に接続される複数のMPEG符号器を含み、前記多重ストリームがビデオ・データの異なるソース、または多重ストリームに分割されるビデオ・データの共通ソースを含む、前記(33)記載のシステム。

(44) ビデオ・フレームの多重ストリームを処理するシステムであって、少なくとも1つの制御可能な符号化パラメータを使用し、1つのストリームを符号化することにより、前記多重ストリームを並列に符号化する複数の符号器と、前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出する手段と、前記多重ストリームの各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記各ストリームの前記符号化を動的に適応化する手段であって、前記動的に適応化する手段が、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整する手段を含み、前記符号化する手段が、前記多重ストリームから獲得される前記少なくとも1つの特性

の相対変化に動的に適応する、適応化手段とを含む、システム。

(45) 前記分析手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出する手段を含む、前記(44)記載のシステム。

(46) 前記複数の符号器が多重圧縮ビデオ・ストリームを生成し、前記システムが、前記圧縮ビデオ・ストリームを多重化し、一定ビット・レート・チャンネル上に転送するためのマルチプレクサを含む、前記(45)記載のシステム。

(47) 前記複数の符号器の1つと前記マルチプレクサとの間に接続される複数のバッファと、前記マルチプレクサと前記一定ビット・レート・チャンネル間に接続されるチャンネル・バッファとを含み、前記分析する手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性に関する統計を導出する手段を含み、前記統計が使用されるビット、平均QUANT、バッファ充填度、またはシーン変化に関する統計の少なくとも1つを含む、前記(46)記載のシステム。

(48) 前記複数の符号器が、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化する手段を含み、前記動的に適応化する手段が、前記各ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットする手段を含む、前記(44)記載のシステム。

(49) 前記動的に適応化する手段が、前記ビデオ・フレームの各ストリームの前記GOPの各々に対して、符号化ビット・レートを決定する手段を含む、前記(48)記載のシステム。

(50) 前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たな前記GOPを開始する手段を含む、前記(49)記載のシステム。

(51) 前記少なくとも1つの特性が、前記ビデオ・フレームの各ストリームに対して、当該ストリームを構成するビデオ・データの複雑度を含み、前記分析手段が、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測する手段を含む、前記(50)記載のシステム。

(52) ビデオ・フレームの多重ストリームの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、前記ビデオ・フレームの多重ストリームを符号化するように、前記コンピュータに指示する手段と、前記多重ストリームを構成する前記ビデオ・フレームの相対複雑度にもとづき、少なくとも1つのストリームの符号化を動的に適応化する

ように、前記コンピュータに指示する手段とを含む、装置。

(53) 前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記少なくとも1つのストリーム内のシーン変化の検出時に、または前記少なくとも1つのストリーム内の新たなフレーム・グループ(GOP)の開始時に、前記少なくとも1つのストリームの符号化の動的な適応化を実行するように、前記コンピュータに指示する手段を含む、前記(52)記載の装置。

(54) 前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記符号化から出力される圧縮ビデオ・ストリームを受信するように接続されるバッファの充填度に関するフィードバックを使用し、前記少なくとも1つのストリームの符号化を動的に適応化するように、前記コンピュータに指示する手段を含む、前記(53)記載の装置。

(55) 前記符号化を動的に適応化するように前記コンピュータに指示する手段が、レート制御アルゴリズムを使用し、前記少なくとも1つのストリームの符号化ビット・レートを制御する手段を含み、前記符号化するように前記コンピュータに指示する手段が、前記多重ストリームの前記符号化から導出される符号化統計を、前記レート制御アルゴリズムに提供するように、前記コンピュータに指示する手段を含む、前記(52)記載の装置。

(56) 前記符号化するように前記コンピュータに指示する手段が、前記多重ストリームを並列に符号化し、多重圧縮ビデオ・ストリームを生成するように、前記コンピュータに指示する手段を含み、前記装置が前記多重圧縮ビデオ・ストリームを一定ビット・レート・チャンネル上に多重化し、転送するように、前記コンピュータに指示する手段を含む、前記(52)記載の装置。

(57) ビデオ・フレームの多重ストリームの処理をコンピュータに指示するコンピュータ読出し可能プログラム・コード手段を有する、少なくとも1つのコンピュータ使用可能媒体を含む装置であって、前記コンピュータ読出し可能プログラム・コード手段が、ビデオ・フレームの各ストリームに対して、少なくとも1つの制御可能な符号化パラメータを使用し、前記多重ストリームを並列に符号化するように、前記コンピュータに指示する手段と、前記各ストリームを分析し、当該ストリームのイントラフレーム特性またはインタフレーム特性のいずれかを含む、少なくとも1つの特性に関する情報を導出するように、前記コンピュータに指示する手段と、各ストリームから獲得される前記少なくとも1つの特性に関する相対情報を使用し、前記ストリームの前記符号化を動的に適応化するように、前記コンピュータに指示する手段であって、前記動的に適応化する手段が、前記各ストリームに対して、前記符号化において使用される前記少なくとも1つの制御可能な符号化パラメータを調整する手段を含み、前記符号化手段が、前記多重ストリームか

ら獲得される前記少なくとも1つの特性の相対変化に動的に適応する、適応化手段とを含む、装置。

(58) 前記符号化を動的に適応化するように前記コンピュータに指示する手段が、前記ビデオ・フレームの各ストリームを分析し、前記少なくとも1つの特性を含む、符号化されるシーンの複雑度に関する情報を導出するように、前記コンピュータに指示する手段を含む、前記(57)記載の装置。

(59) 前記符号化するように前記コンピュータに指示する手段が、前記ビデオ・フレームの各ストリームをフレーム・グループ(GOP)として符号化するように、前記コンピュータに指示する手段を含み、前記動的に適応化する手段が、前記各ストリームの前記GOPの各々の符号化において使用される、前記少なくとも1つの制御可能な符号化パラメータを、前記各GOPの始めに動的にセットするように、前記コンピュータに指示する手段を含む、前記(57)記載の装置。

(60) 前記ビデオ・フレームの各ストリームに対して、当該ストリーム内のシーン変化の検出に際して、新たなフレーム・グループを開始するように前記コンピュータに指示する手段を含み、前記少なくとも1つの特性が、前記各ストリームのビデオ・フレームの複雑度を含み、前記分析するように前記コンピュータに指示する手段が、新たなシーンの最初のフレームの複雑度にもとづき、前記シーンの複雑度を予測するように前記コンピュータに指示する手段を含む、前記(59)記載の装置。

【図面の簡単な説明】

【図1】本発明の原理に従う統計的多重化システムで使用される一般的なMPEG-2準拠の符号器11のフロー図である。

【図2】Iピクチャ、Pピクチャ及びBピクチャの表示順序及び伝送順序、及び前方動き予測及び後方動き予測の例を示す図である。

【図3】本発明に従い実現される多重ストリーム・ビデオ圧縮統計的多重化システムの例を示す図である。

【図4】本発明に従うシーン変化におけるピクチャ・グループ構造の変化を示す図である。

【図5】本発明に従う統計的多重化システムのチャンネル・バッファ充填度対時間の関係を示す図である。

【図6】本発明に従う統計的多重化システムにおいて、

ビデオ・ソースの第1のセットを16Mビット/秒のチャンネル・ビット・レートで符号化するための、ビット・レートの例を示す図である。

【図7】図6のビデオ・ソースの第1のセット($R_c = 16$ Mビット/秒)を符号化するための合計ビット・レートのグラフである。

【図8】図6のビデオ・ソースの第1のセットを、16Mビット/秒のチャンネル・バッファを用い、32Mビット/秒のチャンネル・ビット・レートで記録するための、チャンネル・バッファ充填度のグラフである。

【図9】本発明に従う統計的多重化システムにおいて、ビデオ・ソースの第2のセットを24Mビット/秒のチャンネル・ビット・レートで符号化するための、ビット・レートの例を示す図である。

【図10】図9のビデオ・ソースの第2のセット($R_c = 24$ Mビット/秒)を符号化するための合計ビット・レートを示す図である。

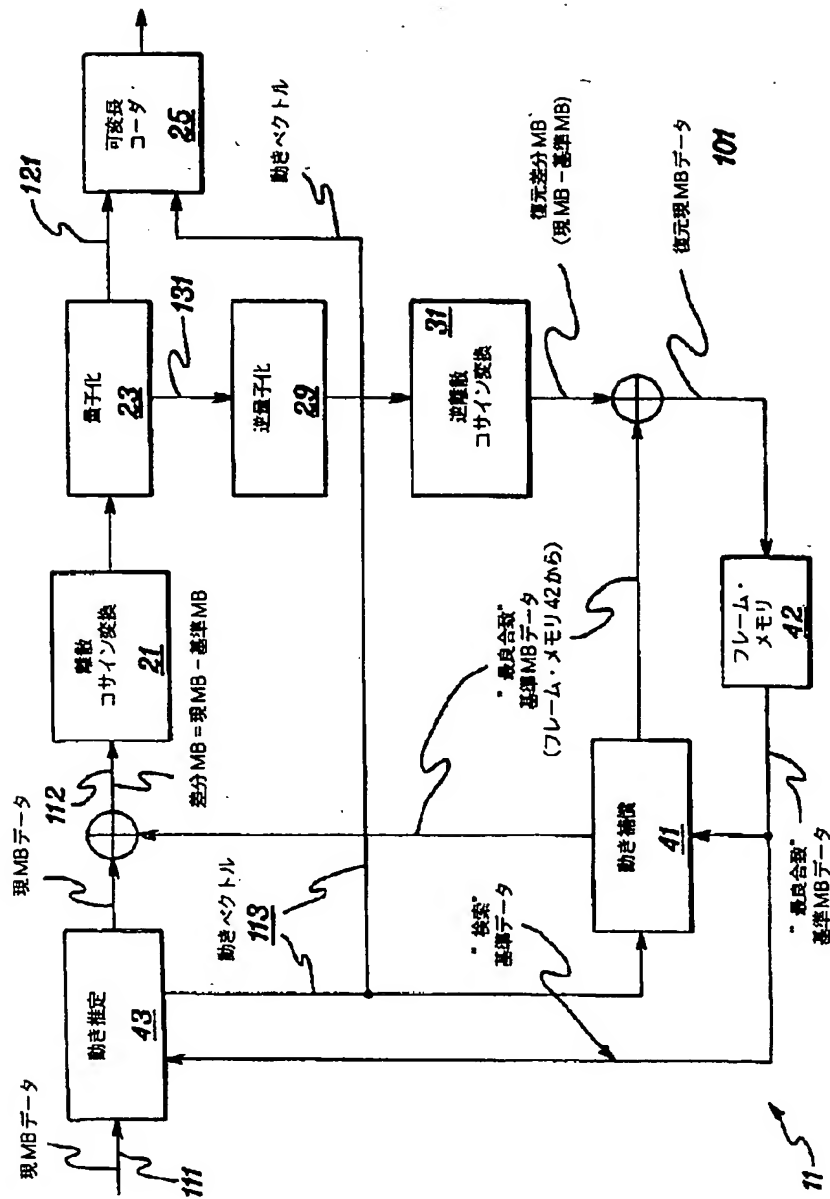
【図11】図9のビデオ・ソースの第2のセットを、12Mビット/秒のチャンネル・バッファを用い、24Mビット/秒のチャンネル・ビット・レートで符号化するための、チャンネル・バッファ充填度のグラフである。

【図12】本発明の原理に従う統計的多重化レート制御処理の1実施例のフロー図である。

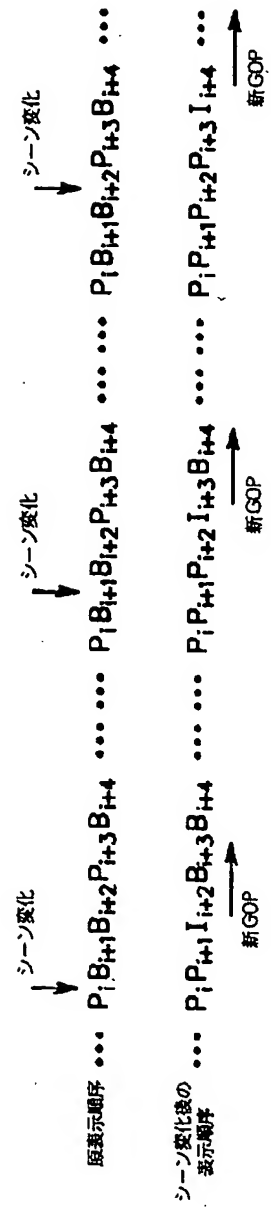
【符号の説明】

- 1 乃至 4、11 符号器
- 21 離散コサイン変換器(DCT)
- 23 量子化器
- 29 逆量子化
- 31 逆離散コサイン変換
- 41 動き補償ユニット
- 42 フレーム・メモリ
- 43 動き推定ユニット
- 113 動きベクトル
- 200 多重番組ビデオ圧縮システム
- 210 MPEG-2ビデオ符号器
- 220 バッファ
- 230 結合レート制御装置
- 240 マルチプレクサ
- 250 チャンネル・バッファ
- 255 チャンネル・バッファ制御フィードバック

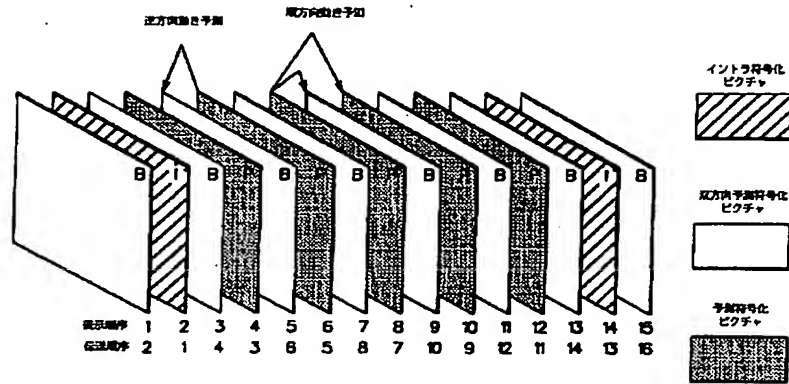
【図1】



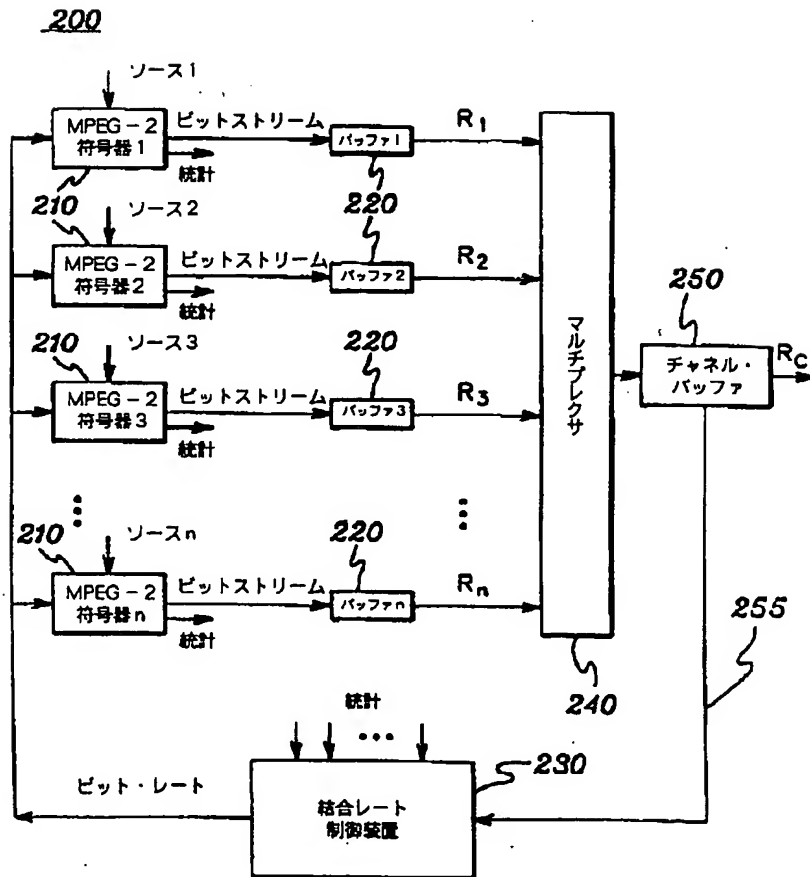
【図4】



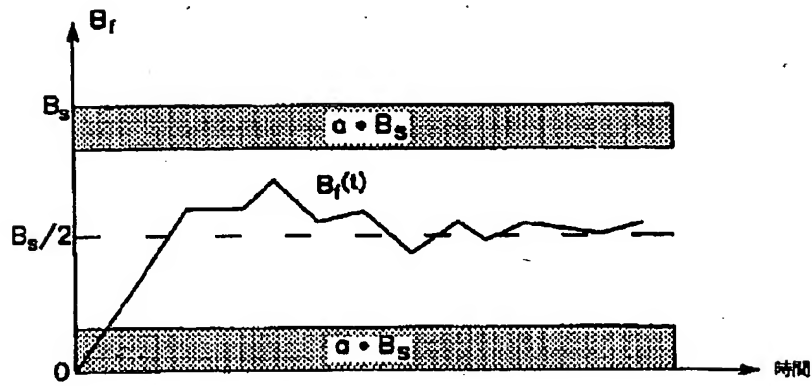
【図2】



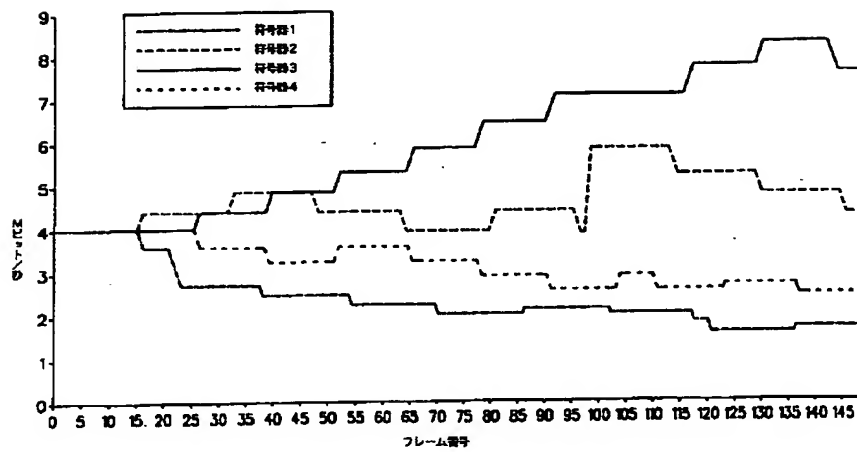
【図3】



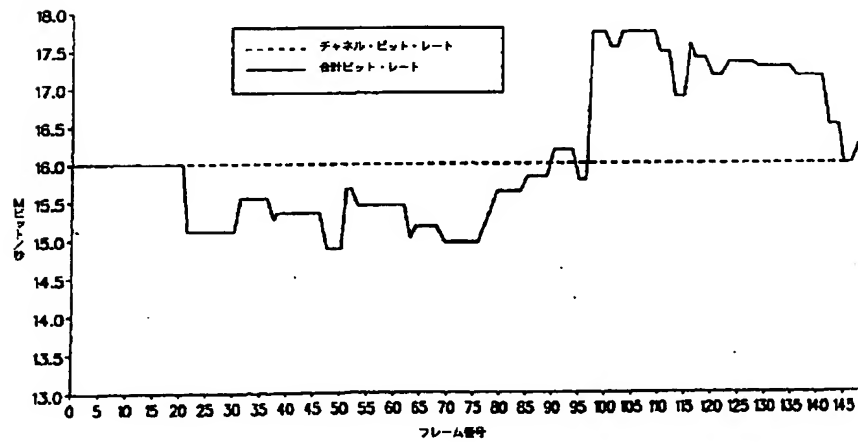
【図5】



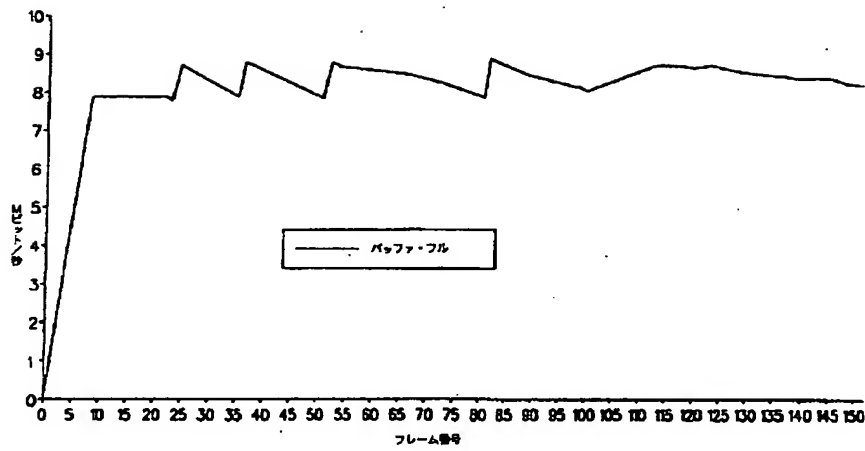
【図6】



【図7】



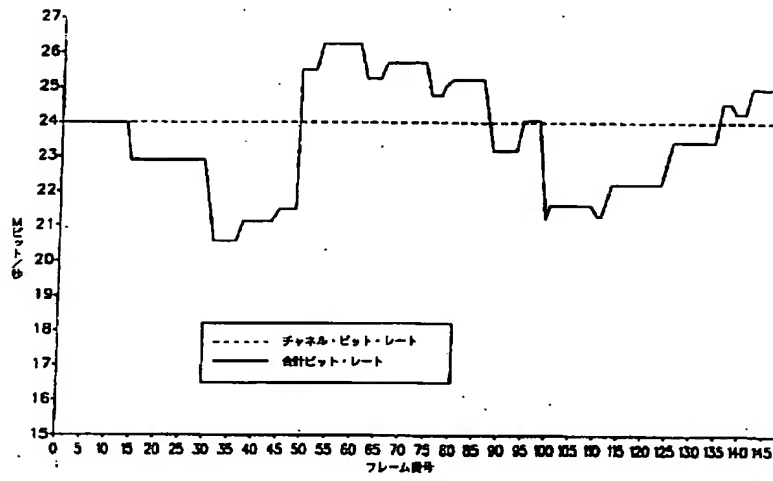
【図8】



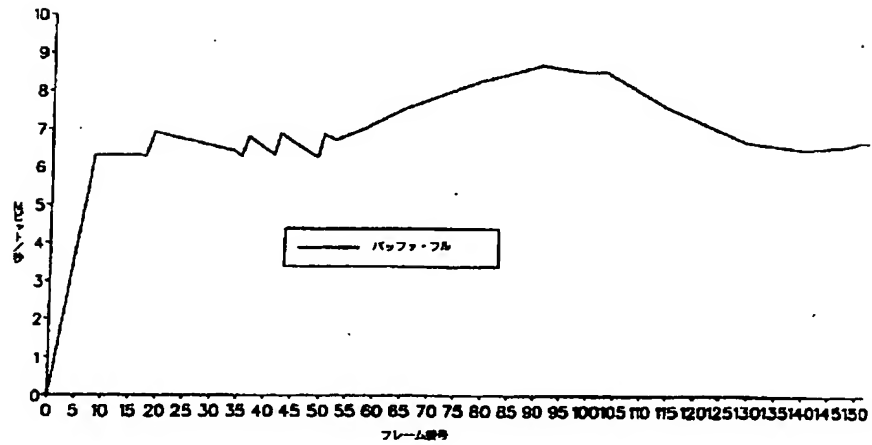
【図9】



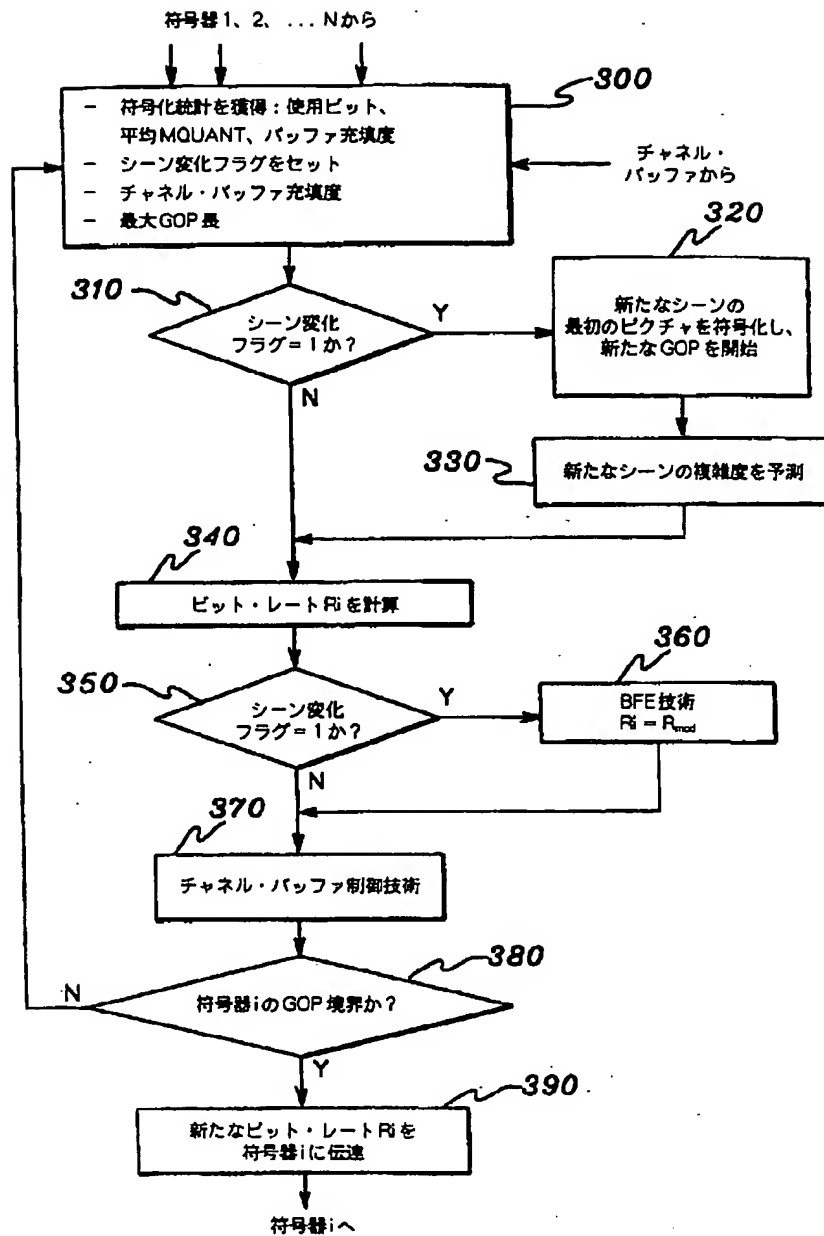
【図10】



【図11】



〔図12〕



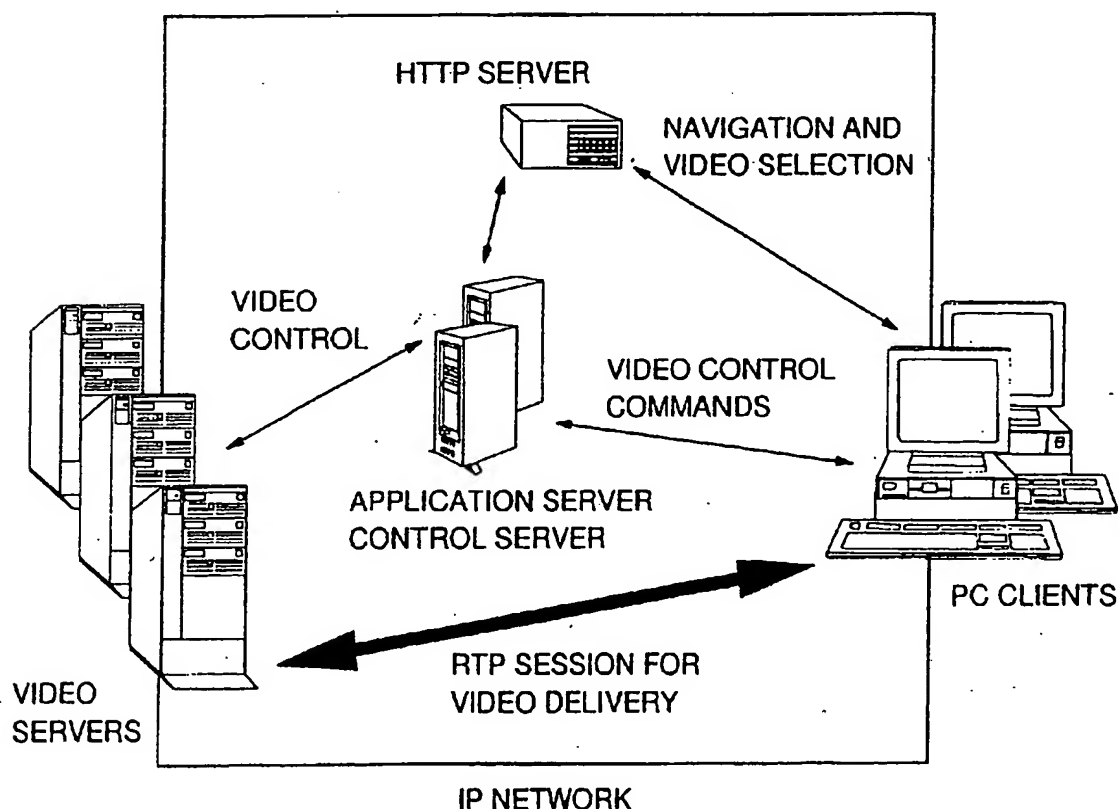
フロントページの続き

(72)発明者 リラ・ボロクスキ
 アメリカ合衆国13760、ニューヨーク州エ
 ンディコット、シエリ・ドライブ 603
 ハーフ

(72)発明者 アグネス・ワイ・ナイ
 アメリカ合衆国13760、ニューヨーク州エ
 ンドウエル、パートリッジ・プレイス
 725

(72)発明者 エドワード・エフ・ウエスターマン
アメリカ合衆国13760、ニューヨーク州エ
ンディコット、ドロシー・ストリート
189

Structured Metadata for Application Specific Viewers for Streamed Internet Video/Audio



Figure

In the system solution shown in the Figure, navigation and selection for video/audio are accomplished using web browser technology with an Hyper Text Transfer Protocol (HTTP) server. In this Figure, other Client-Server types could be used for navigation and selection by changing the HTTP server and appropriate client support (for example; Digital Library client to Digital Library Search server). Video/audio selection can be accomplished in many application scenarios, from selection from a simple list to multimedia objects being integrated into the documents implementing an application.

As a function of selection in this design, metadata is returned from the Application Server to the Client via the HTTP (or other) Server. This metadata causes a helper application or

video/audio viewer to be started in the client. The metadata is passed to this helper application or viewer and is used to initialize the client subsystem for the playback, to get connected to the application server for control functions, and to get the requested title(s) loaded and started. Types of information in the metadata include: address of Application Server, the identifier (unique title) for the file to be played, type of encoding of the Video/Audio, bit rate of the encoded video/audio, etc. Using the receive metadata, all interaction to request and control the play of video is accomplished between the client and the Application Server over the established link. Opening the session will result in a separate logical connection between the Video Server and the client which carries the digitized video. Video Cassette Recorder (VCR)-like control of the video is then supported over the client to Application Server link.

There are classes of viewer applications which will provide very specific functions while the time sensitive data (video/audio) is playing. Examples include text and/or graphics synchronized with the video/audio playback. It is envisioned that there will be application specific viewers designed to perform various types of functions. There are three general functions needed to support the design and operation of these application specific viewers:

1. The ability of the end user or an automated process to select a specific viewer program to be executed for presentation of the requested video/audio. The selection of the viewer program could be based on application and/or user preference. Within the overall process, this viewer selection could be made well in advance of the initiation of the viewer execution: for example, as a function of the selection during browsing or as a function of the Common Gateway Interface (CGI) program in the HTTP Server.
2. The ability of the end user, application designer, or automated process to select/generate other data elements required for execution of the viewer. For example, files (pointers) to the text file to be used for text overlay. Within the overall process, this other data selection could be made well in advance of the initiation of the viewer execution: for example, as a function of the browser selection or as a function of the CGI program in the HTTP Server.
3. The ability for the application designers to select where and how the above decisions and data element generation is accomplished and then the ability to pass this information to the process or viewer to effect the desired result.

There are specific points within the overall selection and application (viewer) initialization process where making these decisions and generation of the appropriate information elements would be useful for an application designer:

1. An HTML program is used for navigation and the selection of the video/audio. Information relative to the Hyper Text Markup Language (HTML) hyperlink selection is sent to the HTTP server. This selection includes the CGI program to be executed in the HTTP Server and information elements which will be passed to this CGI program as parameters. These information elements could be hardcoded into the Universal Resource Locator (URL) of the selection hyperlink, or they could be a function of an HTML form giving the end user the ability to establish the input in various ways. Other Client-Server interaction (e.g., Digital Library) could provide similar functionality.
2. The CGI program, which is executed in the HTTP server as a result of the HTML selection, can use parameters from the selection and/or other information to modify or add to the HTML selection. For example, the CGI program could use a user profile database to effect

the selection of a viewer based on the user preference. Or the CGI program could use a generic parameter from the HTML selection to generate the pointer to the required files (e.g., text overlay) required for this specific selection and viewer. Programs in other server types (e.g., Digital Library Search server) could provide similar functionality.

3. The mime type (file type) of the metadata (from HTTP server to Client) can cause a helper application to be executed in the client endstation. This helper application has access to the information in the metadata file and can interface with the end user. Thus, it could further refine the requests prior to initiation of a specific viewer. For example, it could allow the user input to select a specific viewer.
4. The viewer could provide a user interface which allows the end user to interact with the viewer to influence its execution and function.

With the ability of the customer (or third party) to design and develop application specific viewers, there is a need for them to have access to the information elements passed by the Application Server. This is necessary so they may design the viewers using the appropriate information from the metadata with the calls to the multimedia subsystem interface: establish a session with the Application Server, initiate play of the video/audio file, etc.

One additional problem is to allow for future changes to the Application Server generated metadata (additional functionality) while providing levels of upward and downward operational compatibility between viewers and metadata.

In the system solution shown in the Figure, the metadata was generated exclusive by the Application Server for use by special viewers written to use this data. This metadata was generated and forwarded through the HTTP server as a response to a CGI program request. The metadata could be handled directly by the CGI program (and forwarded to the client), or the CGI program could be notified when the Application Server had completed sending its metadata (via the HTTP Server). In either case, the CGI program has the opportunity to modify and/or append other information elements to the metadata as a mechanism to pass other information to the helper application and/or viewer. By providing a structure for the metadata, other information can be passed from multiple processes to the viewer application and/or viewer unambiguously. The metadata will be structured as follows:

1. The beginning and ending of the Application Server metadata will be marked with keywords so that its bounds can readily be identified.
2. Within the Application Server metadata, the beginning and ending of the Session specific data segments will be marked with keywords so that their bounds can readily be identified. Within the Application Server metadata, the beginning and ending of each selected title specific data segment will be marked with keywords so that bounds of each can readily be identified. The ending of a metadata segment may be indicated by the keyword of the beginning of the next segment.
3. Each field within the Application Server metadata will be encoded as Keyword = Parameter. Thus, the type of field will be self-identifying and the passed parameters easily located. This will allow customers to easily design/implement application specific viewers using the metadata provided by the Application Server. This structure will also allow for new fields to

be added for additional function while maintaining some level of upward/downward compatibility between viewers and Application Server metadata.

4. The CGI program in the HTTP server (or other server type), which requests the metadata generation of the Application Server, will have the opportunity to modify and/or append other information to this metadata file. The CGI program appended information elements could be in front of the Application Server generated metadata and/or behind it.

The capability of the CGI programs to modify/append information to the Application Server generated metadata, allows for the information to be passed from various processes to the helper and/or viewer applications.

1. As a function of authoring the application, specific parameters can be included (hardcoded) in the URL of the HTML document. When the hyperlink is selected, this URL would be passed to the indicated HTTP server and the indicated CGI program initiated. This CGI program can now include URL parameters into the metadata which will now be passed back to the helper and/or viewer application in the client.
2. As a function of authoring the application, forms can be designed in the HTML document which allow the user to make selections. These user selections will become parameters of a URL and can be inspected and/or modified by the CGI program and returned in the metadata file to the helper or viewer application in the client.
3. The initiated CGI program can generate information elements based on its own logic and database information. For example, the CGI program could use a user profile database to set the selection of a user preferred viewer to be used for the application. This information would be passed in the metadata extensions back to the helper application which would then initiated the execution of the appropriate viewer.
4. The above capabilities can be extended to other client-server types (e.g., Digital Library).

This invention disclosure describes the following distinguishing features:

1. A structure for Application Server generated metadata which allows customers to more easily design multimedia viewers.
2. A structure for Application Server generated metadata which allows for upward/downward compatibility between viewers and metadata as future additional functions (information elements) are added to the metadata.
3. A structure for the Application Server generated metadata which allows for the calling server applications (e.g., CGI programs) to modify and append information elements to this metadata. This added information is passed to and used by application specific helper and viewer applications to implement specific functions. The application designer has much flexibility in establishing the origin of these additional information elements within the selection process. The following are possible scenarios for the generation and use of the additional metadata information elements to meet application requirements:
 - The information could be imbedded in the URL of an HTML browser document. This is sent to the HTTP server and becomes parameters to the requested CGI program. Thus, the application designer can dictate parameters to be passed within the authored HTML document. These parameters would normally be transparent (not seen) to the

Structured Metadata for Application Specific Viewers for Streamed Internet Video/Audio — Continued

user. After the CGI program requests the metadata generation from the Application Server, these parameters would be appended to the metadata.

- The information could be generated via an HTML form allowing the end user to make selections. When the form is submitted, these selections are sent to the HTTP server and become parameters to the requested CGI program. The CGI program would then append these to the Application Server generated metadata. Thus, user generated information may be passed to the viewer or helper applications by completing and sending an HTML form.
- The CGI program can use information from other sources (e.g., user profile database) to generate information to be passed to viewer or helper application.
- A helper application used in front of a viewer application may either use appended metadata or may append metadata for the viewer application. A helper application could use metadata information elements to select a specific viewer application to be used for the video playback. A helper application could be used as a front end to a viewer and interact with the user to generate metadata information to be used by the viewer: for example, specific files for the viewer to use for synchronized text/graphics with the video/audio.
- Although the above scenarios were described using HTTP operation, the scenarios could equally apply to other client-server types.